

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 5, 2004, 10:59:47 ; Search time 31.6667 Seconds  
(without alignments)  
89.674 Million cell updates/sec

Title: US-09-761-636A-10  
Perfect score: 50  
Sequence: 1 CSVPLTVC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 840

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*  
1: sp archaea.\*  
2: sp bacteria.\*  
3: sp fungi.\*  
4: sp human.\*  
5: sp invertebrate.\*  
6: sp mammal.\*  
7: sp mhc.\*  
8: sp organelle.\*  
9: sp phage.\*  
10: sp plant.\*  
11: sp rodent.\*  
12: sp virus.\*  
13: sp vertebrate.\*  
14: sp unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriaip.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	40.0	8	2	Q8GM5	Q8gm5 acinetobact
2	18	36.0	9	10	Q9FXL0	Q9fxl0 lilium long
3	17	34.0	7	13	Q8J20	Q8j20 gallus gall
4	17	34.0	7	13	O42564	O42564 fuqu rubrip
5	17	34.0	8	15	Q85562	Q85562 moloney mur
6	17	34.0	9	11	Q8CG13	Q8cg13 mus musculu
7	16	32.0	8	2	Q8RPX4	Q8kpx4 microcystis
8	16	32.0	8	2	Q849P4	Q849p4 salmonella
9	16	32.0	8	4	Q9Y4X6	Q9y4x6 homo sapien
10	16	32.0	8	4	Q16468	Q16468 homo sapien
11	16	32.0	8	10	Q81802	Q81802 zea mays (m
12	16	32.0	9	6	Q9TRU7	Q9tru7 bos taurus
13	16	32.0	9	11	O35953	O35953 mus musculu
14	15	30.0	8	2	Q56140	Q56140 streptococc
15	15	30.0	8	8	Q9TKES	Q9tke5 leptospermu
16	15	30.0	8	8	Q9MD43	Q9md43 rattus norv

17	15	30.0	8	10	Q9SAY7	Q9say7 dioscorea t
18	15	30.0	9	8	Q9TFK2	Q9tkf2 asteromyrtu
19	15	30.0	9	8	Q9TKG1	Q9tkg1 calotobamhus
20	14	28.0	7	2	P70804	P70804 azotobacter
21	14	28.0	7	15	Q07624	Q07624 rous sarcom
22	14	28.0	9	13	Q9PS68	Q9ps68 gallus gall
23	13	26.0	7	12	O67113	O67113 influenzavi
24	13	26.0	8	4	Q15890	Q15890 homo sapien
25	13	26.0	8	4	Q15898	Q15898 homo sapien
26	13	26.0	9	2	O31363	O31363 borrelia ga
27	13	26.0	9	2	Q9K4M6	Q9k4m6 staphylococ
28	13	26.0	9	4	Q15999	Q15999 homo sapien
29	13	26.0	9	8	Q9TKD9	Q9tkd9 pericalymma
30	13	26.0	9	11	Q99JF4	Q99jf4 mus musculu
31	13	26.0	9	12	Q90350	Q90350 hepatitis g
32	13	26.0	9	12	Q9E1U7	Q9e1u7 hepatitis b
33	13	26.0	9	12	O71069	O71069 canine dist
34	13	26.0	9	15	O64972	O64972 avian rous-
35	12	24.0	5	13	P83308	P83308 gallus gall
36	12	24.0	7	10	P93233	P93233 lycopersico
37	12	24.0	7	10	Q9C5B3	Q9c5b3 arabidopsis
38	12	24.0	7	11	O55184	O55184 rattus norv
39	12	24.0	8	2	Q45615	Q45615 bacillus su
40	12	24.0	8	2	O32560	O32560 escherichia
41	12	24.0	8	2	O9X3K1	O9x3k1 prochloroco
42	12	24.0	8	2	P83152	P83152 anabaena sp
43	12	24.0	8	2	O7X4Q5	O7x4q5 nodularia s
44	12	24.0	8	4	Q81V87	Q81v87 homo sapien
45	12	24.0	8	5	O02032	O02032 lytechinus

## ALIGNMENTS

## RESULT 1

Q8GM5 PRELIMINARY; PRT; 8 AA.  
AC Q8GM5;  
DT 01-MAR-2003 (TREMREL. 23, Created)  
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)  
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)  
DE Transposase (Fragment).  
GN TNP17.  
OS Acinetobacter sp. BW3.  
OG Plasmid pKLH207.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Moraxellaceae; Acinetobacter.  
OX NCBI\_TaxID=106395;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BM3; PLASMID=pKLH207;  
RA Kholodii G.Y., Yurieva O.V., Mindlin S.Z., Gorlenko Z.M.,  
RA Nikiforov V.G.;  
RT "pKLH2-like aberrant transposons and possible mechanisms of their  
RT dissemination."  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BM3; PLASMID=pKLH207;  
RA Kholodii G.Y., Mindlin S.Z., Gorlenko Z.M., Yurieva O.V.,  
RA Petrova M.A., Nikiforov V.G.;  
RT "A young family of transposable adaptive DNA segments identified in  
RT the Acinetobacter genus."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ250234; CAC80784.1; -.  
DR EMBL; AJ486856; CAD31078.1; -.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid.  
FT NON\_TER 1  
SQ SEQUENCE 8 AA; 911 MW; 2D71B2D6C1A73774 CRC64;  
Query Match 40.0%; Score 20; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1e+06;

```

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PLTSV 8
Db 2 PLTVQ 6

RESULT 2
Q9FXL0
ID Q9FXL0 PRELIMINARY; PRT; 9 AA.
AC Q9FXL0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE LIM8 protein (Fragment).
GN LIM8.
OS Lilium longiflorum (Trumpet lily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lilium.
OX NCBI_TaxID=4690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hinomoto;
RA Uefuji H., Takase H., Hiratsuka K.;
RT "Lilium longiflorum LIM8 gene, promoter region and partial sequence.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB050987; BAB17856.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1021 MW; 6F8BD76685A6C2CB CRC64;

Query Match 36.0%; Score 18; DB 10; Length 9;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLT 6
Db 5 SMPVT 9

RESULT 3
Q8JJ20
ID Q8JJ20 PRELIMINARY; PRT; 7 AA.
AC Q8JJ20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Extracellular fatty acid binding protein (Fragment).
GN EXPABP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Wang Q., Li N., Li H.;
RT "Cloning and sequencing of 3' UTR of EXPABP gene in chicken.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487519; AAL96665.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 780 MW; 72CBIAB2D5BBB70 CRC64;

Query Match 34.0%; Score 17; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSV 3
Db 2 CSV 4

RESULT 4

```

```

042564
ID 042564 PRELIMINARY; PRT; 7 AA.
AC 042564;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN SCN8A.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97442476; PubMed=9295353;
RA Plummer N.W., McBurney M.W., Meisler M.H.;
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT two-domain protein in fetal brain and non-neuronal cells.";
RL J. Biol. Chem. 272:24008-24015(1997).
DR EMBL; U97673; AAB80916.1; -.
DR GO; GO:0005216; F10n channel activity; IEA.
KW Ionic channel.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 730 MW; 75B72EA2C73772A0 CRC64;

Query Match 34.0%; Score 17; DB 13; Length 7;
Best Local Similarity 71.4%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 3 VPLTSVC 9
Db 1 VPL--VC 5

RESULT 5
Q85562
ID Q85562 PRELIMINARY; PRT; 8 AA.
AC Q85562;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Truncated env protein (Fragment).
OS Moloney murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11801;
RN [1]
RP SEQUENCE OF 4-8 FROM N.A.
RX MEDLINE=82196891; PubMed=6281735;
RA Donoghue D.J., Hunter T.;
RT "A generalized method of subcloning DNA fragments by restriction site
RT reconstruction: Application to sequencing the amino-terminal region of
RT the transforming gene of Gazdar murine sarcoma virus.";
RL Nucleic Acids Res. 10:2549-2564(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164305; PubMed=6300424;
RA Donoghue D.J., Hunter T.;
RT "Recombination junctions of variants of Moloney murine sarcoma virus:
RT Generation and divergence of a mammalian transforming gene.";
RL J. Virol. 45:607-617(1983).
DR EMBL; K03105; AAA46490.1; -.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 732 MW; 98C2D5BEB44DC76D CRC64;

Query Match 34.0%; Score 17; DB 15; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSV 3
Db 5 CSV 7

```

```

RESULT 6
Q8CG13
ID Q8CG13 PRELIMINARY; PRT; 9 AA.
AC Q8CG13;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Glutamate receptor ionotropic N-methyl D-aspartate-like 1A
DE (Fragment).
GN GRIN1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Wyder K.S., Mohan Raj B.K., Sciorra L.J., Roginski R.S.;
RT "The mouse orthologue of the human ionotropic glutamate receptor-like
RT gene (GRIN1A) maps to mouse chromosome 9."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF462417; AA015648.1; -.
DR EMBL; AF462416; AA015648.1; JOINED.
DR MGD; MGI:107282; Grin1a.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON TER 1 1
FT NON TER 9 9
SQ SEQUENCE 9 AA; 1091 MW; 6A91233EB059C33B CRC64;

Query Match 34.0%; Score 17; DB 11; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSVP 4
|:|
Db 6 CKLP 9

RESULT 7
Q8KXP4
ID Q8KXP4 PRELIMINARY; PRT; 8 AA.
AC Q8KXP4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Phycocyanin alpha subunit (Fragment).
GN PCA.
OS Microcystis sp. T96-1.
OC Bacteria; Cyanobacteria; Chroococcales; Microcystis.
OX NCBI_TaxID=198099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=bloom water sample T96-1;
RA Baker J.A., Entsch B., Neilan B.A., McKay D.B.;
RT "Monitoring changing toxigenicity of a cyanobacterial bloom using
RT molecular methods."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY117046; AAM54719.1; -.
FT NON TER 8 8
SQ SEQUENCE 8 AA; 890 MW; F4DB01A73771A336 CRC64;

Query Match 32.0%; Score 16; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PLT 6
|:|
Db 4 PLT 6

```

```

RESULT 8
Q849P4
ID Q849P4 PRELIMINARY; PRT; 8 AA.
AC Q849P4;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PipB (Fragment).
OS Salmonella derby.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=28144;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9813031;
RA Markham P.F., Amavisit P., Lightfoot D., Browning G.F.;
RT "Variation between pathogenic serovars within Salmonella pathogenicity
RT islands."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144492; AAO49836.1; -.
FT NON TER 8 8
SQ SEQUENCE 8 AA; 861 MW; EFC5BDD451A04766 CRC64;

Query Match 32.0%; Score 16; DB 2; Length 8;
Best Local Similarity 40.0%; Pred. No. 1e+06;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VPLTS 7
|:|
Db 1 MPITN 5

RESULT 9
Q9Y4X6
ID Q9Y4X6 PRELIMINARY; PRT; 8 AA.
AC Q9Y4X6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Nuclear LIM interactor (Fragment).
GN NLI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20108806; PubMed=10640831;
RA Drechsler M., Schumacher V., Friedrich S., Wildhardt G., Giesler S.,
RA Schroth A., Bodem J., Royer-Pokora B.;
RT "Genomic structure, alternative transcripts and chromosome location of
RT the human LIM domain binding protein gene LDB1."
RL Cytogenet. Cell Genet. 87:119-124(1999).
DR EMBL; AJ243097; CAB45408.1; -.
FT NON TER 8 8
SQ SEQUENCE 8 AA; 767 MW; EE6EBDBE862D5B6 CRC64;

Query Match 32.0%; Score 16; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSVP 4
|:|
Db 5 CACP 8

RESULT 10
Q16468
ID Q16468 PRELIMINARY; PRT; 8 AA.
AC Q16468;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

```

```

DE DNA for cosmid cC13-1134 PCR primer 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96435920; PubMed=8838806;
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
RA Anand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CpG islands.";
RL Genomics 32:425-430(1996).
DR EMBL; X88976; CAA61407.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871B6 CRC64;

Query Match 32.0%; Score 16; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PLS 6
DB 4 PLS 6

RESULT 11
Q8L802 PRELIMINARY; PRT; 8 AA.
AC Q8L802;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Fat (Fragment).
GN PAT.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Ronning S.B., Berdal K.G., Vaitilingom M.M., Holst-Jensen A.;
RT "Transformation event-specific quantitative real-time PCR for
RT genetically modified Bt11 maize (Zea mays) and estimation of the
RT impact of exogenous DNA on the limit of quantification.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY123624; AAM89275.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 909 MW; 6046C1B2D77412D7 CRC64;

Query Match 32.0%; Score 16; DB 10; Length 8;
Best Local Similarity 40.0%; Pred. No. 1e+06;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PLS 8
DB 4 PLS 8

RESULT 12
Q9TRU7 PRELIMINARY; PRT; 9 AA.
AC Q9TRU7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GAP-3, GTPase-activating protein (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

```

```

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=92112868; PubMed=1309786;
RA Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,
RA Burgess A.W.;
RT "The purification of a Rap1 GTPase-activating protein from bovine
RT brain cytosol.";
RL J. Biol. Chem. 267:1546-1553(1992).
FT NON_TER 1
SQ SEQUENCE 9 AA; 1063 MW; 89EDA77B47604B5A CRC64;

Query Match 32.0%; Score 16; DB 6; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPLTSV 8
DB 4 IPYPSV 9

RESULT 13
O35953 PRELIMINARY; PRT; 9 AA.
AC O35953;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN SCN8A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R111;
RX MEDLINE=97442476; PubMed=9295353;
RA Plummer N.W., McBurney M.W., Meisler M.H.;
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT two-domain protein in fetal brain and non-neuronal cells.";
RL J. Biol. Chem. 272:24008-24015(1997).
DR EMBL; U97672; AAB80914.1; -.
DR MGD; MGI:103169; Scn8a.
DR GQ; CO:0007628; P:adult walking behavior; IMP.
KW Ionic channel.
FT NON_TER 1
SQ SEQUENCE 9 AA; 898 MW; 22D92865B735B737 CRC64;

Query Match 32.0%; Score 16; DB 11; Length 9;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VPLT 6
DB 1 VPLS 4

RESULT 14
Q56140 PRELIMINARY; PRT; 8 AA.
AC Q56140;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE STP6 protein (Fragment).
GN STP6.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]

```



```

RP SEQUENCE FROM N.A.
RC STRAIN=ST11;
RX MEDLINE=95047254; PubMed=7958782;
RA Constable A., Mollet B.;
RT "Isolation and characterisation of promoter regions from Streptococcus
RT thermophilus.";
RL FEMS Microbiol. Lett. 122:85-90(1994).
DR EMBL; X78210; CAA55045.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 846 MW; ED086772D5B045B6 CRC64;

Query Match 30.0%; Score 15; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SVP 4
   |||
Db 4 SVP 6

RESULT 15
Q9TKES PRELIMINARY; PRT; 8 AA.
AC Q9TKES;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE AtpB (Fragment).
GN AtpB.
OS Leptospermum erubescens.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Myrtales; Myrtaceae; Leptospermum.
OX NCBI_TaxID=106049;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Brien M.M., Quinn C.J., Wilson P.G.;
RT "Molecular Systematics of the Leptospermum Suballiance (Myrtaceae).";
RL Aust. J. Bot. 48:0-0(2000).
DR EMBL; AF184690; AAF03860.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER 8
SQ SEQUENCE 8 AA; 876 MW; ECA1B1B764405056 CRC64;

Query Match 30.0%; Score 15; DB 8; Length 8;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PLTS 7
   |||
Db 5 PPTS 8

Search completed: September 5, 2004, 11:05:42
Job time : 32.6667 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 10:56:37 ; Search time 37.3333 Seconds  
(without alignments)  
68.114 Million cell updates/sec

Title: US-09-761-636A-10  
Perfect score: 50  
Sequence: 1 CSVPLTVC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 231240

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Genesep1980s:\*  
2: Genesep1990s:\*  
3: Genesep2000s:\*  
4: Genesep2001s:\*  
5: Genesep2002s:\*  
6: Genesep2003as:\*  
7: Genesep2003bs:\*  
8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	9	4	AAU04529
2	37	74.0	9	5	ABJ04424
3	32	64.0	9	5	ABG34948
4	30	60.0	7	6	ABJ37356
5	30	60.0	7	6	ABJ37436
6	30	60.0	9	5	ABG05266
7	28	56.0	9	7	ADC44660
8	27	54.0	9	2	AAW45666
9	27	54.0	9	2	AAV46691
10	27	54.0	9	5	ABP54842
11	26	52.0	7	7	ABJ78677
12	26	52.0	9	4	AAU03756
13	26	52.0	9	5	ABP54826
14	26	52.0	9	5	ABP54835
15	26	52.0	9	5	ABP54823
16	26	52.0	9	6	ABR75294
17	25	50.0	7	3	ABR75294
18	25	50.0	8	5	ABU04484
19	25	50.0	9	2	AAV48844
20	25	50.0	9	3	AAV64300
21	25	50.0	9	4	AAU03747
22	25	50.0	9	5	ABJ04620
23	25	50.0	9	5	ABJ04630
24	25	50.0	9	5	ABJ04417
25	25	50.0	9	7	ADC44658

26	25	50.0	9	7	ADP78117	Ade78117 Synthetic
27	25	50.0	9	7	ADP78037	Ade78037 Synthetic
28	25	50.0	9	7	ADP78097	Ade78097 Synthetic
29	25	50.0	9	7	ADP7841	Ade7841 Synthetic
30	24	48.0	6	4	AAE05003	Aae05003 Human rel
31	24	48.0	6	4	AAE06020	Aae06020 Human MUM
32	24	48.0	7	3	AAV84175	Aay84175 Amino aci
33	24	48.0	7	4	AAU04530	Aau04530 VEGF base
34	24	48.0	9	2	AAK36888	Aar36888 Insulin-1
35	24	48.0	9	2	AAK43632	Aar43632 Peptide d
36	24	48.0	9	4	AAU03758	Aau03758 Cyclic pe
37	24	48.0	9	4	AAU03731	Aau03731 Cyclic pe
38	24	48.0	9	4	AAU03770	Aau03770 Cyclic pe
39	24	48.0	9	4	AAU03739	Aau03739 Cyclic pe
40	24	48.0	9	4	AAU03760	Aau03760 Cyclic pe
41	24	48.0	9	5	ABG35053	Abg35053 Endostati
42	24	48.0	9	5	ABG68159	Abg68159 Optimised
43	24	48.0	9	5	ABJ04372	Abj04372 Human uro
44	24	48.0	9	5	ABP54841	Abp54841 Alpha-IIB
45	24	48.0	9	5	ABP54837	Abp54837 Alpha-IIB

## ALIGNMENTS

RESULT 1  
AAU04529  
ID AAU04529 standard; peptide; 9 AA.  
XX AC AAU04529;  
XX XX  
DT 26-SEP-2001 (first entry)  
XX XX  
DE VEGF based monocyclic peptide 7.  
XX XX  
KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
KW diabetic retinopathy; chronic inflammation; cyclic.  
XX XX  
OS Synthetic.  
XX XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1...9 /note= "This bond cyclises the peptide"  
XX XX  
PN WO200152875-A1.  
XX XX  
PD 26-JUL-2001.  
XX XX  
PF 18-JAN-2001; 2001WO-US001533.  
XX XX  
PR 18-JAN-2000; 2000US-0176293P.  
XX XX  
PR 16-MAY-2000; 2000US-0204590P.  
XX XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX XX  
PI Achen MG, Hughes RA, Stacker S, Cendron A;  
XX XX  
DR WPI; 2001-442248/47.  
XX XX  
PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
PT or lymphangiogenesis, is produced by cyclising a peptide loop fragment  
PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
PT residues.  
XX XX  
PS Claim 49; Page 32; 102pp; English.  
XX XX  
CC The sequence represents a monomeric monocyclic peptide of the invention,  
CC whose 3-dimensional structure is modelled on the exposed loop of human  
CC VEGF (vascular endothelial growth factor). The invention relates to a  
CC method of producing a monomeric monocyclic peptide by a measuring beta-  
CC beta carbon separation distances on opposite antiparallel strands of a

CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain). The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 50; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9  
 |||||  
 Db 1 CSVPLTSVC 9

RESULT 2  
 ABJ04424  
 ID ABJ04424 standard; peptide; 9 AA.

XX AC ABJ04424;  
 XX DT 24-OCT-2002 (first entry)  
 XX DE Stem cell (mesenchymal) targeting peptide 13.  
 XX KW BRASIL; targeting peptide; bacterial infection;  
 KW Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;  
 KW inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;  
 KW viral infection; cardiovascular disease; degenerative disease.

OS Unidentified.  
 XX WO200220822-A2.  
 XX PD 14-MAR-2002.  
 XX PF 07-SEP-2001; 2001WO-US028124.  
 XX PR 08-SEP-2000; 2000US-0231266P.  
 XX PR 17-JAN-2001; 2001US-00765101.

XX (TEXA ) UNIV TEXAS SYSTEM.  
 XX PI Arap W, Pasqualini R;  
 XX WPI; 2002-404697/43.

XX Identification of targeting peptides that can be used to treat diseases  
 PT e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis  
 PT of Selective Ligands) method comprises a single differential  
 PT centrifugation step.

XX Example 5; Page 75; 167pp; English.

CC The invention comprises a method (BRASIL - Biopanning and Rapid Analysis  
 CC of Selective Interactive Ligands) to obtain a targeting peptide. The  
 CC BRASIL method of the invention involves: exposing a target to a phage  
 CC display library in a first phase; exposing the first phase to a second  
 CC phase; and separating the phage bound to the target from unbound phage.  
 CC The BRASIL method of the invention allows cell phages to be separated  
 CC from the remaining unbound phage in a single differential centrifugation  
 CC step. When compared to conventional cell panning methods, the BRASIL  
 CC method shows a significant increase in recovery of specific phage and a  
 CC substantial decrease in background. The BRASIL method is useful for  
 CC identifying targeting peptides. The targeting peptides identified by the  
 CC method of the invention are useful for treating disease states, such as:  
 CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune  
 CC disease; bacterial infection; viral infection; cardiovascular disease and  
 CC degenerative disease. The present amino acid sequence represents a  
 CC targeting peptide of the invention

XX Sequence 9 AA;

Query Match 74.0%; Score 37; DB 5; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9  
 |||||  
 Db 1 CSVPSVSSC 9

RESULT 3  
 ABG34948  
 ID ABG34948 standard; peptide; 9 AA.

XX AC ABG34948;

XX DT 15-JUL-2002 (first entry)

XX DE Human bone marrow targeting peptide #20.

XX KW Targeting peptide; cancer; Hodgkin's disease; cytostatic;  
 KW immunosuppressive; anti-inflammatory; antiarthritic; antiviral;  
 KW antiatherosclerotic; antidiabetic; antibacterial; diabetes mellitus;  
 KW inflammatory disease; arthritis; atherosclerosis; cancer;  
 KW autoimmune disease; bacterial infection; viral infection.

XX OS Homo sapiens.

XX WO200220722-A2.

XX PD 14-MAR-2002.

XX PF 07-SEP-2001; 2001WO-US027702.

XX PR 08-SEP-2000; 2000US-0231266P.

XX PR 17-JAN-2001; 2001US-00765101.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX PI Arap W, Pasqualini R;

XX WPI; 2002-383050/41.

XX Identifying targeting peptides useful for treating e.g. diabetes  
 PT mellitus, inflammatory diseases, cancer, or autoimmune diseases,  
 PT comprises exposing a sample to a phage display library and recovering  
 PT phage bound to the sample.

XX Claim 56; Page 207; 298pp; English.

XX This invention relates to a novel method for identifying disease  
 CC targeting peptides. The method comprises exposing a sample from an organ,  
 CC tissue or cell type of interest, to a phage display library and  
 CC recovering phage bound to the sample (the phage expresses targeting  
 CC peptides). The peptides identified by the method of the invention may

CC have cytostatic, immunosuppressive, anti-inflammatory, antiarthritic,  
 CC antiatherosclerotic, antidiabetic, antibacterial and antiviral  
 CC activities. The methods and composition are useful for identifying  
 CC targeting peptides and one or more receptors for a targeting peptide. The  
 CC including gene therapy vectors and fusion proteins, to specific organs,  
 CC tissues, or cell types in subject. The targeting peptide may also be used  
 CC for treating diseases such as diabetes mellitus, inflammatory diseases,  
 CC arthritis, atherosclerosis, cancer, autoimmune diseases, bacterial and  
 CC viral infections and Hodgkin's disease. The present sequence represents a  
 CC targeting peptide of the invention  
 XX  
 SQ Sequence 9 AA;

Query Match 64.0%; Score 32; DB 5; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPLTVC 9  
 |||||  
 Db 1 CSPPLTRWC 9

RESULT 4  
 ABJ37356  
 ID ABJ37356 standard; peptide; 7 AA.  
 AC ABJ37356;  
 XX

DT 08-MAY-2003 (first entry)

DE G-protein coupled receptor peptide region #68.

DE Compound library; microenvironment; G-protein Coupled Receptor; GPCR.  
 KW Unidentified.  
 XX

XX WO2003004147-A2.

XX 16-JAN-2003.

XX 05-JUL-2002; 2002WO-GB003094.

XX 06-JUL-2001; 2001GB-00016570.

XX (BIOF-) BIOFOCUS PLC.

XX Crossley R, Rose VS, Stevens AP;

XX WPI; 2003-221549/21.

XX Producing compound library, by generating biological target model using  
 PT target sequence information, defining microenvironments interacting with  
 PT ligand and motifs interacting with microenvironment, and assembling  
 PT motifs.

PS Disclosure; Fig 3; 39pp; English.

XX The invention relates to a novel method for producing a compound library.  
 CC The novel method involves reducing a biological target into a group of  
 CC one or more amino acids required for interaction with a ligand, to  
 CC generate a model of the biological target, using the model to define a  
 CC microenvironment in the biological target, capable of interacting with the  
 CC ligand, defining motifs which interact with the microenvironment, and  
 CC assembling the motifs to generate a compound library for synthesis. The  
 CC novel method is useful to produce compound libraries for screening  
 CC natural ligands such as peptides and proteins or for producing chemical  
 CC compounds based on drug motifs for screening. This sequence represents a  
 CC peptide of a G-protein Coupled Receptor (GPCR), which relates to the  
 CC novel compound library production method of the invention

XX Sequence 7 AA;

Query Match 60.0%; Score 30; DB 6; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSVPLT 6  
 |||||  
 Db 1 CSLPLT 6

RESULT 5  
 ABJ37436  
 ID ABJ37436 standard; peptide; 7 AA.  
 XX

AC ABJ37436;

DT 08-MAY-2003 (first entry)

DE G-protein coupled receptor endothelin ET-A receptor peptide #68.

DE Compound library; microenvironment; G-protein Coupled Receptor; GPCR.  
 KW Unidentified.  
 XX

XX WO2003004147-A2.

XX 16-JAN-2003.

XX 05-JUL-2002; 2002WO-GB003094.

XX 06-JUL-2001; 2001GB-00016570.

XX (BIOF-) BIOFOCUS PLC.

XX Crossley R, Rose VS, Stevens AP;

XX WPI; 2003-221549/21.

XX Producing compound library, by generating biological target model using  
 PT target sequence information, defining microenvironments interacting with  
 PT ligand and motifs interacting with microenvironment, and assembling  
 PT motifs.

PS Disclosure; Fig 7; 39pp; English.

XX The invention relates to a novel method for producing a compound library.  
 CC The novel method involves reducing a biological target into a group of  
 CC one or more amino acids required for interaction with a ligand, to  
 CC generate a model of the biological target, using the model to define a  
 CC microenvironment in the biological target, capable of interacting with the  
 CC ligand, defining motifs which interact with the microenvironment, and  
 CC assembling the motifs to generate a compound library for synthesis. The  
 CC novel method is useful to produce compound libraries for screening  
 CC natural ligands such as peptides and proteins or for producing chemical  
 CC compounds based on drug motifs for screening. This sequence represents a  
 CC peptide of a G-protein Coupled Receptor (GPCR) including Endothelin ET-A  
 CC receptor, which relates to the novel compound library production method  
 CC of the invention

XX Sequence 7 AA;

Query Match 60.0%; Score 30; DB 6; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSVPLT 6  
 |||||  
 Db 1 CSLPLT 6

RESULT 6  
 ABB05266  
 ID ABB05266 standard; peptide; 9 AA.  
 XX



AAW45666 standard; peptide; 9 AA.  
AAW45666;  
27-AUG-2003 (revised)  
09-JUN-1998 (first entry)  
HBV X 69 peptide with binding affinity for HLA-A3-like molecules.  
HLA molecule; cytotoxic T cell; immunogenic peptide; binding affinity;  
HLA-A3 supermotif; tumour; infection; parasite; CTL; antigen; HIV pol;  
HBV; hepatitis B virus.  
Synthetic.  
Hepatitis B virus.  
WO9733602-A1.  
18-SEP-1997.  
10-MAR-1997; 97WO-US003778.  
11-MAR-1996; 96US-0013113P.  
(CYTE-) CYTEL CORP.  
Sette A, Chestnut RW, Sidney J;  
WPI; 1997-470637/43.  
Inducing cytotoxic T cell response against specific antigen - using immunogenic peptide with binding affinity for HLA-A3-like molecules, to treat or prevent tumours and infections by virus, parasites etc.  
Example 1; Page 37; 79pp; English.  
This sequence represents an immunogenic peptide with binding affinity for HLA-A3-like molecules. A cytotoxic T cell (CTL) response against a particular antigen (Ag) is induced in a patient by contacting a CTL with an immunogenic peptide of 9-15 amino acids which binds to at least two HLA-A3-like molecules with dissociation constant less than 500 nM and induces a cytotoxic T cell response. The immunogenic peptide has a sequence of 9 amino acids, comprising a binding motif, with from the N-to C-termini: primary anchor sites (PAR) at positions 2 (selected from Ala, Leu, Ile, Val, Met, Ser or Thr) and 9 (Arg or Lys) and at least one secondary anchor sites (SAR), i.e. Tyr, Phe or Trp at positions 3, 6 or 7, and/or Pro at position 8. The immunogenic peptides are used in peptide based vaccines and therapeutic compositions, for treating viral, parasitic or fungal diseases or cancer, e.g. prostate cancer, hepatitis B or C, renal or cervical carcinoma, lymphoma, cytomegalovirus infection or condyloma acuminatum. They can also be used to elicit a CTL response in vitro for subsequent return of the cells to the patient, e.g. where the patient does not respond to peptide vaccines or other therapies. Selection of specific residues for PAR and SAR results in higher binding affinity and thus increased immunogenicity. (Updated on 27-AUG-2003 to correct OS field.)  
Sequence 9 AA;  
Query Match 54.0%; Score 27; DB 2; Length 9;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CSVPLTS 7  
Db 1 CALPPTS 7  
RESULT 9  
AAW456691  
ID AAY46691 standard; peptide; 9 AA.  
AC AAY46691;

XX 01-DEC-1999 (first entry)  
XX Immunogenic peptide having a human leukocyte antigen binding motif #1302.  
DE Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
XX immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.  
XX Synthetic.  
OS Homo sapiens.  
XX WO9945954-A1.  
PN 16-SEP-1999.  
XX 13-MAR-1998; 98WO-US005039.  
PF 13-MAR-1998; 98WO-US005039.  
PR (EPIM-) EPIMMUNE INC.  
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
PI WPI; 1999-551214/46.  
DR New immunogenic peptides with HLA binding motif, useful in treatment and diagnosis of cancers and viral diseases.  
PT Claim 1; Page 81; 150pp; English.  
XX AAY45390 to AAY48214 represent specifically claimed immunogenic peptides having a human major histocompatibility complex (MHC) Class I (also known as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against the antigen from which the peptide is derived. Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are normally induced by an antigen in the form of a peptide fragment bound to a HLA molecule, rather than the intact foreign antigen itself, and are particularly important in tumour rejection and in fighting viral infections. The peptides are therefore useful therapeutically to treat or prevent viral infections and cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma. They can be administered as vaccines to elicit an immune response in individuals susceptible or otherwise at risk of viral infection or cancer, or used to treat chronic or acute conditions. They are also useful diagnostically, and can be used to induce a cytotoxic T cell response, by contacting a cytotoxic T cell with the peptide e.g. to produce CTLs ex vivo for infusion back into a patient. The polynucleotides encoding the immunogenic peptides are also useful therapeutically and for immunisation as above  
Sequence 9 AA;  
Query Match 54.0%; Score 27; DB 2; Length 9;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CSVPLTS 7  
Db 1 CALPPTS 7  
RESULT 10  
ABP54842  
ID ABP54842 standard; peptide; 9 AA.  
XX ABP54842;  
AC ABP54842;  
XX 08-JAN-2003 (first entry)  
XX

DE Alpha-IIB beta-3 integrin activating peptide.  
 XX  
 KW Alpha-IIB beta-3 integrin activator; integrin; CD41; CD61;  
 KW glycoprotein IIB/IIIA; von Willebrand disease; coagulant; vulnerary;  
 KW cyclic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..9  
 FT Region 3..5  
 FT /note= "binding motif, region specifically described in  
 FT Claim 1"  
 FT  
 XX WO200272619-A1.  
 XX  
 XX 19-SEP-2002.  
 XX  
 XX 12-MAR-2002; 2002WO-FI000193.  
 XX  
 XX 12-MAR-2001; 2001FI-00000492.  
 XX  
 XX (CTTC-) CTT CANCER TARGETING TECHNOLOGIES OY.  
 XX  
 XX Koivunen E, Gahmberg CG;  
 XX  
 XX WPI; 2002-750482/81.  
 XX  
 XX New alphaIIB beta3 integrin activating peptides useful for manufacturing  
 PT a composition for treating or preventing thrombotic or bleeding  
 PT disorders, e.g. von Willebrand disease, and in wound healing and tissue  
 PT regeneration.  
 XX  
 XX Disclosure; Page 12; 34pp; English.  
 XX  
 XX The present sequence is that of a novel alpha-IIB beta-3 integrin  
 CC (glycoprotein IIB/IIIA or CD41/CD61) activating peptide. This cyclic  
 CC peptide comprises a consensus binding motif (Val-Pro-Trp) and was  
 CC identified by focusing a phage library screening on integrin ligands  
 CC which are not blocked by a GRGDS peptide. The VWP motif is present in the  
 CC A3-domain of von Willebrand factor (vWF), suggesting that vWF is an  
 CC activator of the alpha-IIB beta-3 complex, enabling stable platelet-vWF  
 CC interaction. The invention relates to the use of novel peptides (see  
 CC AP54823-25) comprising the consensus binding motif as pharmaceuticals  
 CC for the treatment of thrombotic diseases and bleeding disorders,  
 CC including von Willebrand disease, and in artificial tissue transplants to  
 CC aid in wound healing and tissue regeneration (all claimed)  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 54.0%; Score 27; DB 5; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CSVPLTSVC 9  
 DB 1 CDVFWRLDC 9  
 RESULT 11  
 ADB79677  
 ID ADB79677 standard; peptide; 7 AA.  
 XX  
 AC ADB79677;  
 XX  
 XX 04-DEC-2003 (first entry)  
 DT  
 XX Parapoxvirus ORF 100 C-terminal peptide.  
 DE  
 XX virucide; anti-HIV; hepatotropic; antiinflammatory; cytostatic;  
 KW vulnerary; antiasthmatic; antiallergic; dermatological; antidiabetic;  
 KW immunosuppressive; antirheumatic; antiarthritic; thyromimetic;  
 KW protozoacide; amoebicide; antibacterial; gene therapy; virus;

KW viral infections; non-viral infections; proliferative disease;  
 KW inflammatory disease; allergic disease; autoimmune disease.  
 XX Parapoxvirus.  
 OS  
 XX WO2003006654-A2.  
 XX  
 XX 23-JAN-2003.  
 XX  
 XX 12-JUN-2002; 2002WO-EP006440.  
 XX  
 XX 13-JUN-2001; 2001NZ-00512341.  
 XX  
 XX (FARB ) BAYER AG.  
 XX  
 XX Weber O, Friederichs SM, Siegling A, Schlapp T, Mercer AA;  
 PI Fleming SB;  
 XX  
 XX WPI; 2003-221750/21.  
 XX  
 XX New polynucleotide and recombinant proteins of Parapoxvirus ovis, useful  
 PT for manufacturing a medicament for treating virus related disease, viral  
 PT infections, non-viral infections, proliferative disease or inflammatory  
 PT disease.  
 XX  
 XX Example 4; Page 37; 51pp; English.  
 XX  
 XX The invention relates to a novel purified and isolated polynucleotide  
 CC (N1) of Parapoxvirus ovis (PPVO) comprising a nucleotide sequence (S1,  
 CC not defined in the specification), or its complementary sequence, or  
 CC fragment or functional variant. A polynucleotide of the invention has  
 CC virucide, anti-HIV, hepatotropic, antiinflammatory, cytostatic,  
 CC vulnerary, antiasthmatic, antiallergic, dermatological, antidiabetic,  
 CC immunosuppressive, antirheumatic, antiarthritic, thyromimetic.  
 CC CC protozoacide, amoebicide, and antibacterial activity. The polynucleotides  
 CC may have a use in gene therapy. The recombinant proteins encoded by the  
 CC polynucleotides, or recombinant viruses comprising a vaccinia virus  
 CC genome and fragments of a PPVO genome are useful for manufacturing  
 CC pharmaceutical compositions for treating virus related disease (e.g.  
 CC hepatitis, papillomatosis, herpes virus infections, liver fibrosis, HIV  
 CC infections or influenza), viral infections, non-viral infections (e.g.  
 CC infections with mycobacteria, mycoplasma, amoeba or plasmodia), skin  
 CC proliferative disease (e.g. cancer, leukaemia, warts or other skin  
 CC neoplasms), inflammatory disease (e.g. Crohn's disease, COPD, asthma or  
 CC conditions related to healing of wounds), allergic disease, and/or  
 CC autoimmune diseases (systemic lupus erythematosus, Sjogren's disease,  
 CC Hashimoto's thyroiditis, rheumatoid arthritis or diabetes mellitus). The  
 CC present sequence is used in the exemplification of the invention.  
 XX  
 XX Sequence 7 AA;  
 SQ  
 Query Match 52.0%; Score 26; DB 7; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 PLTSVC 9  
 DB 1 PLTGM 6  
 RESULT 12  
 AAU03756  
 ID AAU03756 standard; peptide; 9 AA.  
 XX  
 XX AAU03756;  
 AC  
 XX 26-SEP-2001 (first entry)  
 DT  
 XX Cyclic peptide inhibitor of LFA-1/ICAM-1 interaction #33.  
 DE  
 XX Cyclic; lymphocyte function associated antigen-1; LFA-1; asthma;  
 KW intracellular adhesion molecule; ICAM-1; inhibitor; leukaemia;  
 KW haematopoietic neoplastic disease; myocardial infarction;



KW radiation injury; rheumatoid arthritis; lymphoma metastasis;  
 KW retinoic acid syndrome; all-trans retinoic acid.  
 OS Synthetic.

PN WO200151508-A1.  
 PD 19-JUL-2001.

PF 16-JAN-2001; 2001WO-US001382.  
 XX 14-JAN-2000; 2000US-00483550.  
 PR 16-JAN-2001; 2001US-00760599.

XX (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
 PA Larson RS;  
 PI WPI; 2001-432906/46.

DR Composition comprising a cyclic peptide inhibitor of lymphocyte function  
 XX associated antigen-1 and intracellular adhesion molecule 1 interaction,  
 PT for treating e.g. asthma and myocardial infarction.

PS Example 2; Page 23; 58pp; English.  
 XX The sequence represents the amino acid sequence of cyclic peptide  
 CC inhibitor #33 of lymphocyte function associated antigen-1 and  
 CC intracellular adhesion molecule (LFA-1/ICAM-1) interaction. A composition  
 CC comprising a cyclic peptide inhibitor of LFA-1/ICAM-1 interaction is  
 CC useful for treating haematopoietic neoplastic disease, myocardial  
 CC infarction, radiation injury, asthma, rheumatoid arthritis or lymphoma  
 CC metastasis. The composition is also useful for inhibiting in a subject  
 CC the interaction between LFA-1 expressed on a leukocyte and ICAM-1  
 CC expressed on another cell, preventing retinoic acid syndrome in a subject  
 CC receiving all-trans retinoic acid, inhibiting growth of leukaemia cells,  
 CC inhibiting emigration of leukocytes from blood into tissue and screening  
 CC a candidate compound for binding to ICAM-1

XX Sequence 9 AA;  
 SQ Query Match 52.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+06;  
 Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPPLTSVC 9  
 DB 1 CALMRSSIC 9

RESULT 13  
 ABP54826  
 ID ABP54826 standard; peptide; 9 AA.

XX ABP54826;  
 AC 08-JAN-2003 (first entry)

DE Alpha-IIB beta-3 integrin activating peptide.  
 XX Alpha-IIB beta-3 integrin activator; integrin; CD41; CD61;  
 KW glycoprotein IIB/IIIA; von Willebrand disease; coagulant; vulnery;

KW cyclic.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 FH Disulfide-bond 1..9  
 FT Region 3..5  
 FT /note= "binding motif, region specifically described in  
 FT Claim 1"

PN WO200272619-A1.

XX 19-SEP-2002.  
 PD 12-MAR-2002; 2002WO-FI000193.  
 PF 12-MAR-2001; 2001FI-00000492.  
 PR (CTTC-) CTT CANCER TARGETING TECHNOLOGIES OY.

PA Koivunen E, Gahmberg CG;  
 PI WPI; 2002-750482/81.  
 DR New alphaIIB beta3 integrin activating peptides useful for manufacturing  
 XX a composition for treating or preventing thrombotic or bleeding  
 PT disorders, e.g. von Willebrand disease, and in wound healing and tissue  
 FT regeneration.

PS Disclosure; Page 13; 34pp; English.  
 XX The present sequence is that of a novel alpha-IIB beta-3 integrin  
 CC (glycoprotein IIB/IIIA or CD41/CD61) activating peptide. This cyclic  
 CC peptide comprises a consensus binding motif (Val-Pro-Trp) that was  
 CC identified by focusing a phage library screening on integrin ligands,  
 CC which are not blocked by a GRGDS peptide. The motif is present in the A3-  
 CC domain of von Willebrand factor (vWf), suggesting that vWf is an  
 CC activator of the alpha-IIB beta-3 complex, enabling stable platelet-vWf  
 CC interaction. The invention relates to the use of novel peptides  
 CC comprising the consensus binding motif as pharmaceuticals for the  
 CC treatment of thrombotic diseases and bleeding disorders, including von  
 CC Willebrand disease, and in artificial tissue transplants to aid in wound  
 CC healing and tissue regeneration (all claimed)

XX Sequence 9 AA;  
 SQ Query Match 52.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CSVPPLTSVC 9  
 DB 1 CAVEPWARYC 9

RESULT 14  
 ABP54835  
 ID ABP54835 standard; peptide; 9 AA.

XX ABP54835;  
 AC 08-JAN-2003 (first entry)

DE Alpha-IIB beta-3 integrin activating peptide.  
 XX Alpha-IIB beta-3 integrin activator; integrin; CD41; CD61;  
 KW glycoprotein IIB/IIIA; von Willebrand disease; coagulant; vulnery;

KW cyclic.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 FH Disulfide-bond 1..9  
 FT Region 3..5  
 FT /note= "binding motif, region specifically described in  
 FT Claim 1"

PN WO200272619-A1.  
 XX 19-SEP-2002.  
 PF 12-MAR-2002; 2002WO-FI000193.  
 PR 12-MAR-2001; 2001FI-00000492.

XX	(CTTC-) CTT CANCER TARGETING TECHNOLOGIES OY.	12-MAR-2001; 2001FI-00000492.
PA	Koivunen E, Gahmberg CG;	
PI	WPI; 2002-750482/81.	
DR		
XX	New alphaIIb beta3 integrin activating peptides useful for manufacturing	
PT	a composition for treating or preventing thrombotic or bleeding	
PT	disorders, e.g. von Willebrand disease, and in wound healing and tissue	
PT	regeneration.	
XX		
PS	Disclosure; Page 12; 34pp; English.	
XX		
CC	The present sequence is that of a novel alpha-IIb beta-3 integrin	
CC	(Glycoprotein IIb/IIIa or CD41/CD61) activating peptide. This cyclic	
CC	peptide comprises a consensus binding motif (Val-Pro-Trp) and was	
CC	identified by focusing a phage library screening on integrin ligands	
CC	which are not blocked by a GRGDS peptide. The VW motif is present in the	
CC	A3-domain of von Willebrand factor (vWf), suggesting that vWf is an	
CC	activator of the alpha-IIb beta-3 complex, enabling stable platelet-vwf	
CC	interaction. The invention relates to the use of novel peptides (see	
CC	ABP54823-25) comprising the consensus binding motif as pharmaceuticals	
CC	for the treatment of thrombotic diseases and bleeding disorders,	
CC	including von Willebrand disease, and in artificial tissue transplants to	
CC	aid in wound healing and tissue regeneration (all claimed)	
XX		
SQ	Sequence 9 AA;	
	Query Match 52.0%; Score 26; DB 5; Length 9;	
	Best Local Similarity 44.4%; Pred. No. 1.4e+06;	
	Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY	1 CSVPLTSVC 9	
	:       :	
DB	1 CAVPWGRLC 9	
	:       :	
RESULT 15		
ABP54823		
ID	ABP54823 standard; peptide; 9 AA.	
XX		

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:05:54 ; Search time 35.6667 Seconds  
(without alignments)  
79.502 Million cell updates/sec

Title: US-09-761-636A-10  
Perfect score: 50  
Sequence: 1 CSVPLTSVC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 119143

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PTC\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PTCUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	9 US-09-761-636A-10	Sequence 10, Appl
2	32	64.0	9	12 US-10-363-208-24	Sequence 24, Appl
3	30	60.0	9	9 US-09-832-723-98	Sequence 98, Appl
4	30	60.0	9	14 US-10-303-331-98	Sequence 98, Appl
5	28	56.0	9	14 US-10-286-457-389	Sequence 389, Appl
6	26	52.0	9	9 US-09-760-599-34	Sequence 34, Appl
7	26	52.0	9	14 US-10-254-446A-147	Sequence 147, Appl
8	25	50.0	8	12 US-10-462-452-480	Sequence 480, Appl
9	25	50.0	8	16 US-10-601-953-609	Sequence 609, Appl
10	25	50.0	8	16 US-10-322-266-481	Sequence 481, Appl
11	25	50.0	9	9 US-09-760-599-25	Sequence 25, Appl
12	25	50.0	9	14 US-10-006-869-3614	Sequence 3614, Ap
13	25	50.0	9	14 US-10-286-457-387	Sequence 387, Appl
14	25	50.0	9	15 US-10-395-032-3614	Sequence 3614, Ap
15	24	48.0	7	9 US-09-761-636A-11	Sequence 11, Appl

16	24	48.0	7	12	US-10-458-334-2	Sequence 2, Appli
17	24	48.0	9	8	US-08-344-824-293	Sequence 293, App
18	24	48.0	9	9	US-09-760-599-9	Sequence 9, Appli
19	24	48.0	9	9	US-09-760-599-17	Sequence 17, Appl
20	24	48.0	9	9	US-09-760-599-36	Sequence 36, Appl
21	24	48.0	9	9	US-09-760-599-38	Sequence 38, Appl
22	24	48.0	9	9	US-09-760-599-48	Sequence 48, Appl
23	24	48.0	9	10	US-09-747-802-6	Sequence 6, Appli
24	24	48.0	9	10	US-09-932-165-284	Sequence 284, App
25	24	48.0	9	12	US-10-363-208-140	Sequence 140, App
26	24	48.0	9	12	US-09-935-430-249	Sequence 249, App
27	24	48.0	9	12	US-09-935-430-322	Sequence 322, App
28	24	48.0	9	12	US-09-935-430-349	Sequence 349, App
29	24	48.0	9	12	US-09-935-430-464	Sequence 464, App
30	24	48.0	9	14	US-10-277-292-249	Sequence 249, App
31	24	48.0	9	14	US-10-277-292-322	Sequence 322, App
32	24	48.0	9	14	US-10-277-292-349	Sequence 349, App
33	24	48.0	9	14	US-10-277-292-464	Sequence 464, App
34	24	48.0	9	15	US-10-280-340-249	Sequence 249, App
35	24	48.0	9	15	US-10-280-340-322	Sequence 322, App
36	24	48.0	9	15	US-10-280-340-349	Sequence 349, App
37	24	48.0	9	15	US-10-280-340-464	Sequence 464, App
38	24	48.0	9	15	US-10-107-532-250	Sequence 250, App
39	24	48.0	9	15	US-10-107-532-256	Sequence 256, App
40	24	48.0	9	15	US-10-107-532-257	Sequence 257, App
41	24	48.0	9	15	US-10-107-532-778	Sequence 778, App
42	24	48.0	9	15	US-10-107-532-792	Sequence 792, App
43	24	48.0	9	15	US-10-107-532-809	Sequence 809, App
44	24	48.0	9	15	US-10-107-532-1314	Sequence 1314, Ap
45	24	48.0	9	15	US-10-107-532-1319	Sequence 1319, Ap

## ALIGNMENTS

RESULT 1  
US-09-761-636A-10  
; Sequence 10, Application US/09761636A  
; Patent No. US20020085218A1  
; GENERAL INFORMATION:  
; APPLICANT: ACHEN, Marc  
; APPLICANT: STACKER, Steven  
; APPLICANT: HUGHES, Richard  
; APPLICANT: GENDRON, Angela  
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR  
; FILE REFERENCE: 1064/48505 Achen et al  
; CURRENT APPLICATION NUMBER: US/09/761,636A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/176,293  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/204,590  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-636A-10

Query Match 100.0% Score 50; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSVPLTSVC 9  
| | | | | | | | | |  
Db 1 CSVPLTSVC 9

RESULT 2  
US-10-363-208-24  
; Sequence 24, Application US/10363208  
; Publication No. US20040048243A1

```

; GENERAL INFORMATION:
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: Methods and Compositions for In Vitro Targeting
; FILE REFERENCE: 005774.P005PCT
; CURRENT APPLICATION NUMBER: US/10/363,208
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(9)
; OTHER INFORMATION: synthetic construct
US-10-363-208-24

Query Match          64.0%; Score 32; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CSVPPLTSVC 9
      || || || || |
Db      1 CSPPLTRWC 9

RESULT 3
US-09-832-723-98
; Sequence 98, Application US/09832723
; Patent No. US20020098524A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; APPLICANT: Chen, Yiyou
; APPLICANT: Murray, Christopher J.
; APPLICANT: Tijerina, Pilar
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
; FILE REFERENCE: GC617-2
; CURRENT APPLICATION NUMBER: US/09/832,723
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,259
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptides screened from a phage display random
; OTHER INFORMATION: peptide library
US-09-832-723-98

Query Match          60.0%; Score 30; DB 9; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 CSVPPLTSVC 9
      || || || || |
Db      1 CKMPTSKVC 9

RESULT 4
US-10-303-331-98
; Sequence 98, Application US/10303331
; Publication No. US20030152976A1
; GENERAL INFORMATION:
; APPLICANT: Janssen, Giselle G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Winetzk, Deborah S.
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
; FILE REFERENCE: GC617-3
; CURRENT APPLICATION NUMBER: US/10/303,331

```

```

; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: US 09/832,723
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,259
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptides screened from a phage display random
; OTHER INFORMATION: peptide library
US-10-303-331-98

Query Match          60.0%; Score 30; DB 14; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 CSVPPLTSVC 9
      || || || || |
Db      1 CKMPTSKVC 9

RESULT 5
US-10-286-457-389
; Sequence 389, Application US/10286457
; Publication No. US20030166004A1
; GENERAL INFORMATION:
; APPLICANT: JENO GYURIS et al.
; TITLE OF INVENTION: ENDOTHELIAL-CELL BINDING PEPTIDES FOR DIAGNOSIS AND THERAPY
; FILE REFERENCE: GPC1-P01-178
; CURRENT APPLICATION NUMBER: US/10/286,457
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/334822
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 684
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 389
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence isolated from random peptide libraries, based
; OTHER INFORMATION: ability to selectively bind to endothelial cells
US-10-286-457-389

Query Match          56.0%; Score 28; DB 14; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.2e+06;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 CSVPPLTSVC 9
      || || || || |
Db      1 CDLPTRSRIC 9

RESULT 6
US-09-760-599-34
; Sequence 34, Application US/09760599
; Patent No. US20010034326A1
; GENERAL INFORMATION:
; APPLICANT: Larson Mr., Richard S.
; TITLE OF INVENTION: Peptide Inhibitors of LFA-1/CAM-1 Interaction
; FILE REFERENCE: SCI200/4-1C1P
; CURRENT APPLICATION NUMBER: US/09/760,599
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic  
US-09-760-599-34

Query Match 52.0%; Score 26; DB 9; Length 9;  
Best Local Similarity 33.3%; Pred. No. 1.2e+06;  
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9  
| : : | : |  
Db 1 CALWESIC 9

RESULT 7  
US-10-254-446A-147  
; Sequence 147, Application US/10254446A  
; Publication No. US20030113714A1  
; GENERAL INFORMATION:  
; APPLICANT: Belcher, Angela M  
; APPLICANT: Smalley, Richard E.  
; APPLICANT: Ryan, Esther  
; APPLICANT: Lee, Seung-Wuk  
; TITLE OF INVENTION: BIOLOGICAL CONTROL OF NANOPARTICLES  
; FILE REFERENCE: 02-02US  
; CURRENT APPLICATION NUMBER: US/10/254,446A  
; CURRENT FILING DATE: 2003-02-19  
; PRIOR APPLICATION NUMBER: 60/325,664  
; PRIOR FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 245  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 147  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: peptide with peptide binding sequence retrieved from phage biopar

US-10-254-446A-147

Query Match 52.0%; Score 26; DB 14; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.2e+06;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9  
| : : | : |  
Db 1 CKLQLTNQC 9

RESULT 8  
US-10-462-452-480  
; Sequence 480, Application US/10462452  
; Publication No. US20040037809A1  
; GENERAL INFORMATION:  
; APPLICANT: Quay, Steven  
; APPLICANT: El Shary, Mohammed Abd  
; APPLICANT: Gupta, Malini  
; APPLICANT: de Meireles, Jorge  
; TITLE OF INVENTION: Compositions and Methods for Enhanced  
; FILE REFERENCE: 02-02US  
; CURRENT APPLICATION NUMBER: US/10/462,452  
; CURRENT FILING DATE: 2003-06-16  
; PRIOR APPLICATION NUMBER: 60/393,066  
; PRIOR FILING DATE: 2002-06-28  
; NUMBER OF SEQ ID NOS: 790  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 480  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-462-452-480

Query Match 50.0%; Score 25; DB 12; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PLTSVC 9  
| : : | : |  
Db 1 PVTVC 6

RESULT 9  
US-10-601-953-609  
; Sequence 609, Application US/10601953  
; Publication No. US2004007540A1  
; GENERAL INFORMATION:  
; APPLICANT: Quay, Steven C.  
; TITLE OF INVENTION: Compositions And Methods For Modulating Physiology Of Epithelial  
; FILE REFERENCE: 02-03US  
; CURRENT APPLICATION NUMBER: US/10/601,953  
; CURRENT FILING DATE: 2003-06-24  
; PRIOR APPLICATION NUMBER: 60/392,512  
; PRIOR FILING DATE: 2002-06-28  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 609  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic construct  
US-10-601-953-609

Query Match 50.0%; Score 25; DB 16; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PLTSVC 9  
| : : | : |  
Db 1 PVTVC 6

RESULT 10  
US-10-322-266-481  
; Sequence 481, Application US/10322266  
; Publication No. US20040115135A1  
; GENERAL INFORMATION:  
; APPLICANT: Quay, Steven C.  
; TITLE OF INVENTION: Compositions And Methods For Enhanced Mucosal Delivery Of Peptide  
; FILE REFERENCE: NPC10567  
; CURRENT APPLICATION NUMBER: US/10/322,266  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 797  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 481  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic construct  
US-10-322-266-481

Query Match 50.0%; Score 25; DB 16; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PLTSVC 9  
| : : | : |  
Db 1 PVTVC 6

RESULT 11  
US-09-760-599-25  
; Sequence 25, Application US/09760599

```
; Patent No. US20010034326A1
; GENERAL INFORMATION:
; APPLICANT: Larson Mr., Richard S.
; TITLE OF INVENTION: Peptide Inhibitors of LFA-1/CAM-1 Interaction
; FILE REFERENCE: SCI200/4-1CIP
; CURRENT APPLICATION NUMBER: US/09/760,599
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-760-599-25

Query Match          50.0%; Score 25; DB 9; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.2e+06;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1  CSVPLTSVC 9
DB      1  CMLRMNSIC 9

RESULT 12
US-10-006-869-3614
; Sequence 3614, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3614
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
; OTHER INFORMATION: recognition sequence
US-10-006-869-3614

Query Match          50.0%; Score 25; DB 14; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1  CSVPLTSVC 9
DB      1  CTFHIDSVC 9

RESULT 13
US-10-286-457-387
; Sequence 387, Application US/10286457
; Publication No. US20030166004A1
; GENERAL INFORMATION:
; APPLICANT: JENO GYURIS et al.
; TITLE OF INVENTION: ENDOTHELIAL-CELL BINDING PEPTIDES FOR DIAGNOSIS AND THERAPY
; FILE REFERENCE: GPCI-P01-178
; CURRENT APPLICATION NUMBER: US/10/286,457
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/334822
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 684
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 387
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence isolated from random peptide libraries, based
; OTHER INFORMATION: ability to selectively bind to endothelial cells
US-10-286-457-387

Query Match          50.0%; Score 25; DB 14; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1  CSVPLTSVC 9
DB      1  CPDPTTRLC 9

RESULT 14
US-10-395-032-3614
; Sequence 3614, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3614
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
; OTHER INFORMATION: recognition sequence
US-10-395-032-3614

Query Match          50.0%; Score 25; DB 15; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1  CSVPLTSVC 9
DB      1  CTFHIDSVC 9

RESULT 15
US-09-761-636A-11
; Sequence 11, Application US/09761636A
; Patent No. US20020065218A1
; GENERAL INFORMATION:
; APPLICANT: ACHEN, Marc
; APPLICANT: STACKER, Steven
; APPLICANT: HUGHES, Richard
; APPLICANT: CENDRON, Angela
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR
; FILE REFERENCE: 1064/48505 Achen et al
; CURRENT APPLICATION NUMBER: US/09/761,636A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/176,293
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: US 60/204,590
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
```

; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-636A-11

Query Match 48.0%; Score 24; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VPLTS 7  
|||  
Db 2 VPLTS 6  
|||

Search completed: September 5, 2004, 11:16:00  
Job time : 36.6667 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:01:38 ; Search time 12 Seconds  
(without alignments)  
38.719 Million cell updates/sec

Title: US-09-761-636A-10

Perfect score: 50

Sequence: 1 CSVPITVSVC 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 89883

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pcp:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pcp:\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pcp:\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pcp:\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pcp:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	52.0	9	4	US-09-760-599-34
2	25	50.0	9	3	US-09-258-754-199
3	25	50.0	9	3	US-09-042-107-199
4	25	50.0	9	4	US-09-187-859-3614
5	25	50.0	9	4	US-09-839-542B-3614
6	25	50.0	9	4	US-09-722-250D-199
7	25	50.0	9	4	US-09-760-599-25
8	25	50.0	9	4	US-09-483-550B-25
9	24	48.0	6	1	US-08-481-434A-21
10	24	48.0	6	3	US-08-476-134A-30
11	24	48.0	6	6	5190920-26
12	24	48.0	6	6	5506208-28
13	24	48.0	9	1	US-07-958-903A-43
14	24	48.0	9	1	US-08-462-018-43
15	24	48.0	9	1	US-08-823-245-43
16	24	48.0	9	2	US-08-598-873-49
17	24	48.0	9	3	US-08-605-430-49
18	24	48.0	9	4	US-07-963-329A-43
19	24	48.0	9	4	US-09-760-599-9
20	24	48.0	9	4	US-09-760-599-17
21	24	48.0	9	4	US-09-760-599-36
22	24	48.0	9	4	US-09-760-599-38
23	24	48.0	9	4	US-09-760-599-48
24	24	48.0	9	4	US-09-483-550B-9
25	24	48.0	9	4	US-09-483-550B-17
26	24	48.0	9	5	PCT-US92-09443A-43
27	23	46.0	8	4	US-08-475-955-67

28	23	46.0	9	1	US-08-195-075-4	Sequence 4, Appli
29	23	46.0	9	1	US-08-467-083-3	Sequence 3, Appli
30	23	46.0	9	1	US-08-414-417B-3	Sequence 3, Appli
31	23	46.0	9	2	US-08-486-348A-3	Sequence 3, Appli
32	23	46.0	9	2	US-08-468-545B-3	Sequence 3, Appli
33	23	46.0	9	2	US-08-986-234-80	Sequence 80, Appli
34	23	46.0	9	3	US-08-466-680B-3	Sequence 3, Appli
35	23	46.0	9	3	US-09-258-754-146	Sequence 146, App
36	23	46.0	9	3	US-09-258-754-308	Sequence 308, App
37	23	46.0	9	3	US-09-139-802-92	Sequence 92, Appli
38	23	46.0	9	3	US-08-660-092-122	Sequence 122, App
39	23	46.0	9	3	US-09-042-107-146	Sequence 146, App
40	23	46.0	9	3	US-09-042-107-308	Sequence 308, App
41	23	46.0	9	4	US-09-160-513-122	Sequence 122, App
42	23	46.0	9	4	US-09-659-786-92	Sequence 92, Appli
43	23	46.0	9	4	US-08-403-459-25	Sequence 25, Appli
44	23	46.0	9	4	US-08-926-914-92	Sequence 92, Appli
45	23	46.0	9	4	US-09-722-250D-146	Sequence 146, App

## ALIGNMENTS

RESULT 1  
US-09-760-599-34  
; Sequence 34, Application US/09760599  
; Patent No. 6630447  
; GENERAL INFORMATION:  
; APPLICANT: Larson Mr., Richard S.  
; TITLE OF INVENTION: Peptide Inhibitors of LFA-1/CAM-1 Interaction  
; FILE REFERENCE: SCT200/4-ICIP  
; CURRENT APPLICATION NUMBER: US/09/760,599  
; CURRENT FILING DATE: 2001-01-17  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-760-599-34

Query Match 52.0%; Score 26; DB 4; Length 9;  
Best Local Similarity 33.3%; Pred. No. 3e+05;  
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPITVSVC 9  
Db 1 CALMRSSIC 9

RESULT 2  
US-09-258-754-199  
; Sequence 199, Application US/09258754  
; Patent No. 6174887  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Rajotte, Daniel  
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
; TITLE OF INVENTION: Membrane Dipeptidase  
; FILE REFERENCE: P-LJ 3443  
; CURRENT APPLICATION NUMBER: US/09/258,754  
; CURRENT FILING DATE: 1999-02-26  
; EARLIER APPLICATION NUMBER: 09/042,107  
; EARLIER FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 452  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 199  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-199

Query Match          50.0%; Score 25; DB 3; Length 9;
Best Local Similarity 55.6%; Pred. No. 3e+05; 4; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CSVPLTSVC 9
Db 1 CSAYTTSPC 9

RESULT 3
US-09-042-107-199
; Sequence 199, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: MOLECULES THAT HOME TO VARIOUS SELECTED ORGANS OR
; TITLE OF INVENTION: TISSUES
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-199

Query Match          50.0%; Score 25; DB 3; Length 9;
Best Local Similarity 55.6%; Pred. No. 3e+05; 4; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CSVPLTSVC 9
Db 1 CSAYTTSPC 9

RESULT 4
US-09-187-859-3614
; Sequence 3614, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3614
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
; OTHER INFORMATION: recognition sequence
US-09-187-859-3614

Query Match          50.0%; Score 25; DB 4; Length 9;
Best Local Similarity 44.4%; Pred. No. 3e+05; 3; Indels 0; Gaps 0;
Matches 4; Conservative 2; Mismatches 2;

QY 1 CSVPLTSVC 9
Db 1 CSAYTTSPC 9

RESULT 5
US-09-839-542B-3614
; Sequence 3614, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3614
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
; OTHER INFORMATION: recognition sequence
US-09-839-542B-3614

Query Match          50.0%; Score 25; DB 4; Length 9;
Best Local Similarity 44.4%; Pred. No. 3e+05; 3; Indels 0; Gaps 0;
Matches 4; Conservative 2; Mismatches 2;

QY 1 CSVPLTSVC 9
Db 1 CTFHIDSVC 9

RESULT 6
US-09-722-250D-199
; Sequence 199, Application US/09722250D
; Patent No. 6610651
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: MOLECULES THAT HOME TO VARIOUS SELECTED ORGANS OR
; TITLE OF INVENTION: TISSUES
; FILE REFERENCE: P-LJ 4514
; CURRENT APPLICATION NUMBER: US/09/722,250D
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 09/042,107
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-722-250D-199

Query Match          50.0%; Score 25; DB 4; Length 9;
Best Local Similarity 55.6%; Pred. No. 3e+05; 4; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CSVPLTSVC 9
Db 1 CSAYTTSPC 9

RESULT 7
US-09-760-599-25

```

```

; Sequence 25, Application US/09760599
; Patent No. 6630447
; GENERAL INFORMATION:
; APPLICANT: Larson Mr., Richard S.
; TITLE OF INVENTION: Peptide Inhibitors of LFA-1/CAM-1 Interaction
; FILE REFERENCE: SCI200/4-ICIP
; CURRENT APPLICATION NUMBER: US/09/760,599
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-760-599-25

Query Match          50.0%; Score 25; DB 4; Length 9;
Best Local Similarity 33.3%; Pred. No. 3e+05;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9
Db 1 CMLRMNSIC 9

RESULT 8
US-09-483-550B-25
; Sequence 25, Application US/09483550B
; Patent No. 6649592
; GENERAL INFORMATION:
; APPLICANT: Larson Mr., Richard S.
; TITLE OF INVENTION: Peptide Inhibitors of LFA-1/CAM-1 Interaction
; FILE REFERENCE: SCI200/4-001
; CURRENT APPLICATION NUMBER: US/09/483,550B
; CURRENT FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-483-550B-25

Query Match          50.0%; Score 25; DB 4; Length 9;
Best Local Similarity 33.3%; Pred. No. 3e+05;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9
Db 1 CMLRMNSIC 9

RESULT 9
US-08-483-434A-21
; Sequence 21, Application US/08483434A
; Patent No. 5648461
; GENERAL INFORMATION:
; APPLICANT: EVAL, Jacob
; APPLICANT: HAMILTON, Bruce K.
; APPLICANT: TUSZYNSKI, George P.
; TITLE OF INVENTION: Synthetic Analogs of Thrombospondin and
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA

```

```

; ZIP: 19103-2398
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,434A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/450,738
; FILING DATE: 25-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/185,614
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/024,436
; FILING DATE: 01-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/587,197
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/483,527
; FILING DATE: 22-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-306 (9049)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 567-2020
; TELEFAX: (215) 567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-434A-21

Query Match          48.0%; Score 24; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSVP 4
Db 1 CSVP 4

RESULT 10
US-08-476-134A-30
; Sequence 30, Application US/08476134A
; Patent No. 6239110
; GENERAL INFORMATION:
; APPLICANT: EYAL, JACOB
; APPLICANT: HAMILTON, BRUCE K.
; APPLICANT: TUSZYNSKI, GEORGE P.
; TITLE OF INVENTION: SYNTHETIC ANALOGS OF THROMBOSPONDIN AND THERAPEUTIC USE
; FILE REFERENCE: 07206-0009
; CURRENT APPLICATION NUMBER: US/08/476,134A
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/587,197
; PRIOR FILING DATE: 1990-09-24
; PRIOR APPLICATION NUMBER: 07/483,527
; PRIOR FILING DATE: 1990-02-22
; PRIOR APPLICATION NUMBER: 08/450,738
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: 08/185,614
; PRIOR FILING DATE: 1994-01-24
; PRIOR APPLICATION NUMBER: 08/024,436

```

Db 1 CSVP 4

RESULT 13

US-07-958-903A-43

; Sequence 43, Application US/07958903A

; Patent No. 5652214

; GENERAL INFORMATION:

; APPLICANT: Lewis, Michael E.

; APPLICANT: Kauer, James C.

; APPLICANT: Smith, Kevin R.

; APPLICANT: Callison, Kathleen V.

; APPLICANT: Baldino, Frank

; APPLICANT: Neff, Nicola

; APPLICANT: Iqbal, Mohamed

; TITLE OF INVENTION: TREATING DISORDERS BY APPLICATION

; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH FACTORS AND

; TITLE OF INVENTION: ANALOGS

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM PS/2 Model 50Z or 55SX

; OPERATING SYSTEM: MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/958,903A

; FILING DATE: October 7, 1992

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/361,595

; FILING DATE: June 5, 1989

; APPLICATION NUMBER: 07/534,139

; FILING DATE: June 5, 1990

; APPLICATION NUMBER: 07/869,913

; FILING DATE: April 15, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 02655/003004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-8906

; TELEFAX: (617) 542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; US-07-958-903A-43

Query Match 48.0%; Score 24; DB 1; Length 9;

Best Local Similarity 44.4%; Pred. No. 3e+05; Indels 0;

Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 CSVPLTSVC 9

Db 1 CCTPAKSEC 9

RESULT 14

US-08-462-018-43

; Sequence 43, Application US/08462018

; Patent No. 5703045

; GENERAL INFORMATION:

; APPLICANT: Lewis, Michael E.

; APPLICANT: Kauer, James C.  
 ; APPLICANT: Smith, Kevin R.  
 ; APPLICANT: Callison, Kathleen V.  
 ; APPLICANT: Baldino, Frank  
 ; APPLICANT: Neff, Nicola  
 ; APPLICANT: Iqbal, Mohamed  
 ; TITLE OF INVENTION: TREATING DISORDERS BY APPLICATION  
 ; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH FACTORS AND  
 ; TITLE OF INVENTION: ANALOGS  
 ; NUMBER OF SEQUENCES: 56  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM PS/2 Model 502 or 55SX  
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)  
 ; SOFTWARE: WordPerfect (Version 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/462,018  
 ; FILING DATE:

; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/958,903  
 ; FILING DATE: October 7, 1992  
 ; APPLICATION NUMBER: 07/361,595  
 ; FILING DATE: June 5, 1989  
 ; APPLICATION NUMBER: 07/534,139  
 ; FILING DATE: June 5, 1990  
 ; APPLICATION NUMBER: 07/869,913  
 ; FILING DATE: April 15, 1992  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.  
 ; REGISTRATION NUMBER: 30,162  
 ; REFERENCE/DOCKET NUMBER: 02655/003005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 43:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9  
 ; TYPE: amino acid  
 ; STRANDEDNESS: linear  
 ; TOPOLOGY: linear  
 ; US-08-462-018-43

Query Match 48.0%; Score 24; DB 1; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSVPLTSVC 9  
 Db 1 CCTPAKSEC 9

## RESULT 15

US-08-823-245-43  
 ; Sequence 43, Application US/08823245  
 ; Patent No. 5776897

; GENERAL INFORMATION:  
 ; APPLICANT: Lewis, Michael  
 ; APPLICANT: Kauer, James C.  
 ; APPLICANT: Smith, Kevin R.  
 ; APPLICANT: Callison, Kathleen V.  
 ; APPLICANT: Baldino, Frank  
 ; APPLICANT: Neff, Nicola  
 ; APPLICANT: Iqbal, Mohamed  
 ; TITLE OF INVENTION: TREATING DISORDERS BY

; TITLE OF INVENTION: APPLICATION  
 ; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH  
 ; TITLE OF INVENTION: FACTORS AND  
 ; TITLE OF INVENTION: ANALOGS  
 ; NUMBER OF SEQUENCES: 56  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM PS/2 Model 502 or 55SX  
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)  
 ; SOFTWARE: WordPerfect (Version 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/823,245  
 ; FILING DATE: March 24, 1997

; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/361,595  
 ; FILING DATE: June 6, 1989  
 ; APPLICATION NUMBER: 07/534,139  
 ; FILING DATE: June 5, 1990  
 ; APPLICATION NUMBER: 07/869,913  
 ; FILING DATE: April 15, 1992  
 ; APPLICATION NUMBER: 07/958,903  
 ; FILING DATE: October 7, 1992

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Creeson, Gary L.  
 ; REGISTRATION NUMBER: 34,310  
 ; REFERENCE/DOCKET NUMBER: 02655/003008  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 43:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9  
 ; TYPE: amino acid  
 ; STRANDEDNESS: N/A  
 ; TOPOLOGY: N/A

US-08-823-245-43

Query Match 48.0%; Score 24; DB 1; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 3e+05;  
 Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSVPLTSVC 9  
 Db 1 CCTPAKSEC 9

Search completed: September 5, 2004, 11:07:05  
 Job time : 13 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:14:14 ; Search time 21 Seconds  
(without alignments)  
32.064 Million cell updates/sec

Title: US-09-761-636A-11

Perfect score: 42

Sequence: 1 CVP L T S C 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	47.6	5	2	B22565	R-phycoerythrin al
2	15	35.7	6	2	H48394	Glycoprotein compo
3	15	35.7	6	2	I65546	MHC H2-L antigen -
4	15	35.7	7	2	B34818	vicillin 57K chain
5	15	35.7	7	2	A34026	acetylcholinestera
6	14	33.3	5	2	E60274	major protein anti
7	13	31.0	7	2	S42620	aggreccan - bovine
8	12	28.6	6	2	C22565	R-phycoerythrin be
9	12	28.6	6	2	I67345	MHC H2-K-k cell su
10	12	28.6	7	2	S08606	hypothetical prote
11	11	26.2	4	2	A20393	tyrosine-melanocyt
12	11	26.2	4	2	I54357	schwannomin - mous
13	11	26.2	5	2	A60521	glycogen phosphory
14	11	26.2	6	2	I49421	laminin B1 - weste
15	11	26.2	7	2	E61491	seed protein ws-5
16	11	26.2	7	2	I48105	dihydrofolate redu
17	11	26.2	7	2	I48086	DNA topoisomerase
18	10	23.8	5	2	F22565	R-phycoerythrin ga
19	10	23.8	6	2	I37263	y protein - human
20	10	23.8	7	2	EGMUCR	catch-relaxing pep
21	10	23.8	7	2	A61081	tryptophyllin, bas
22	10	23.8	7	2	S38516	mablinin II chain
23	10	23.8	7	2	PR0087	ribulose-bisphosph
24	10	23.8	7	2	A28340	myomodulin - Calif
25	10	23.8	7	2	E30608	Ig kappa chain V-I
26	10	23.8	7	4	I56695	hypothetical L2 pr
27	10	23.8	7	4	A58725	viotoxin - destro
28	9	21.4	3	3	A22565	R-phycoerythrin al
29	9	21.4	4	2	I51049	metallothionein-A

30 9 21.4 4 2 S43959  
31 9 21.4 4 2 S55238  
32 9 21.4 5 2 A33882  
33 9 21.4 5 2 B37988  
34 9 21.4 5 2 B45525  
35 9 21.4 5 2 S65726  
36 9 21.4 5 2 S11127  
37 9 21.4 5 2 H4817  
38 9 21.4 5 2 F44817  
39 9 21.4 5 2 B44817  
40 9 21.4 5 2 D44817  
41 9 21.4 6 2 A61049  
42 9 21.4 6 2 JU0355  
43 9 21.4 6 2 PQ0008  
44 9 21.4 6 2 A31263  
45 9 21.4 6 2 A27696

## ALIGNMENTS

### RESULT 1

B22565  
R-phycoerythrin alpha-2 chain - red alga (Gastroclonium coulteri) (fragment)  
C:Species: Gastroclonium coulteri  
C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993  
R:Accession: B22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: B22565  
A:Molecule type: protein  
A:Residues: 1-5 <KLO>

Query Match 47.6%; Score 20; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVP 3  
DB 2 CVP 4

### RESULT 2

H48394  
Glycoprotein component 16/major fat-globule membrane protein/WFG-B8 homolog - bovine (fr  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C:Accession: H48394  
R:Mather, I.H.; Banghart, L.R.; Lane, W.S.  
Biochem. Mol. Biol. Int. 29, 545-554, 1993  
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.  
A:Reference number: A48394; MUID:93250576; PMID:8485470  
A:Accession: H48394  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <MAT>  
A:Experimental source: milk  
A>Note: sequence extracted from NCBI backbone (NCBIP:131518)  
C:Keywords: glycoprotein

Query Match 35.7%; Score 15; DB 2; Length 6;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 VPLTSC 7  
DB 1 VELLGC 6

### RESULT 3

I65546  
MHC H2-L antigen - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I65546  
R:Kimura, A.; Israel, A.; Le Bail, O.; Kourileky, P.  
Cell 44, 261-272, 1986  
A:Title: Detailed analysis of the mouse H-2Kb promoter: Enhancer-like sequences and their  
A:Reference number: 152778; MUID:86106202; PMID:3510743  
A:Accession: I65546  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: GB:M12483; NID:gl9565; PIDN:AAA39663.1; PID:g554234

Query Match 35.7%; Score 15; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VPLT 5  
| |  
| |  
Db 2 VPCT 5

RESULT 4  
B34818  
vicillin 57K chain - pigeon pea (fragment)  
C:Species: Cajanus cajan (pigeon pea)  
C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 30-Sep-1993  
C:Accession: B34818  
R:Mawal, Y.R.; Mawal, M.R.; Ranjekar, P.K.  
Biochem. Biophys. Res. Commun. 166, 1446-1452, 1990  
A:Title: Unusual denaturation properties of vicillin from Cajanus cajan.  
A:Reference number: A34818; MUID:90165956; PMID:2306256  
A:Accession: B34818  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MAW>

Query Match 35.7%; Score 15; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TSC 7  
| |  
| |  
Db 1 TTC 3

RESULT 5  
A34026  
acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)  
C:Species: Torpedo californica (Pacific electric ray)  
C:Date: 16-Mar-1990 #sequence\_revision 16-Mar-1990 #text\_change 08-Nov-1996  
C:Accession: A34026  
R:Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.; T  
J. Biol. Chem. 263, 1140-1145, 1988  
A:Title: Divergence in primary structure between the molecular forms of acetylcholinesterase  
A:Reference number: A34026; MUID:88087239; PMID:3335534  
A:Accession: A34026  
A:Molecule type: protein  
A:Residues: 1-7 <GIB>  
C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 35.7%; Score 15; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 TSC 7  
| |  
| |  
Db 5 TAC 7

RESULT 6

E60274  
major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)  
C:Species: Mycobacterium tuberculosis  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 30-Sep-1993  
C:Accession: E60274  
R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.  
Infect. Immun. 59, 372-382, 1991  
A:Title: Isolation and partial characterization of major protein antigens in the culture  
A:Reference number: A60274; MUID:9109989; PMID:1898899  
A:Accession: E60274  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-5 <NAG>

Query Match 33.3%; Score 14; DB 2; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLT 5  
| |  
| |  
Db 3 PIT 5

RESULT 7  
S42620  
aggreccan - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 24-Jul-1998 #text\_change 24-Jul-1998  
C:Accession: S42620  
R:Vogel, K.G.; Sandy, J.D.; Pogany, G.; Robbins, J.R.  
Matrix Biol. 14, 171-179, 1994  
A:Title: Aggreccan in bovine tendon.  
A:Reference number: S42620; MUID:94340214; PMID:7520336  
A:Accession: S42620  
A:Molecule type: protein  
A:Residues: 1-7 <VOG>  
A:Experimental source: flexor tendon  
C:Keywords: cartilage

Query Match 31.0%; Score 13; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PLTS 6  
| |  
| |  
Db 2 PIVS 5

RESULT 8  
C22565  
R-phycoerythrin beta-1 chain - red alga (Gastrocionium coulteri) (fragment)  
C:Species: Gastrocionium coulteri  
C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993  
C:Accession: C22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: C22565  
A:Molecule type: protein  
A:Residues: 1-6 <KLO>

Query Match 28.6%; Score 12; DB 2; Length 6;  
Best Local Similarity 25.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTSC 7  
| |  
| |  
Db 1 MAAC 4

RESULT 9  
167345



MHC H2-K-k cell surface glycoprotein - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C:Accession: I67345  
 R:Archibald, A.L.; Thompson, N.A.; Kvist, S.  
 EMBO J. 5, 957-965, 1986  
 A:Title: A single nucleotide difference at the 3' end of an intron causes differential splicing of the H2-K-k gene  
 A:Reference number: I53243; MUID:86247587; PMID:3013627  
 A:Accession: I67345  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-6 <RES>  
 A:Cross-references: GB:M26859; NID:9199439; PIDN:AAA39612.1; PID:G387458  
 C:Genetics: 6/1  
 A:Introns: 6/1  
 C:Keywords: glycoprotein

Query Match 28.6%; Score 12; DB 2; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 LTSC 7  
 DB 1 LPDC 4

RESULT 10  
 S08606  
 hypothetical protein 2 estrogen receptor 5'-region - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 18-Aug-2000  
 C:Accession: S08606  
 R:Krust, A.; Green, S.; Argos, P.; Kumar, P.; Bornert, J.M.; Chambon, P.  
 EMBO J. 5, 891-897, 1986  
 A:Title: The chicken estrogen receptor sequence: homology with v-erbA and the human oestrogen receptor  
 A:Reference number: S07192; MUID:86247578; PMID:3755102  
 A:Accession: S08606  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-7 <XRU>  
 A:Cross-references: EMBL:X03805; NID:G53378; PIDN:CAA27432.1; PID:G584490  
 C:Superfamily: unassigned leader peptides

Query Match 28.6%; Score 12; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 LTSC 7  
 DB 3 LAHC 6

RESULT 11  
 A32039  
 tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 31-Jul-1989 #sequence\_revision 31-Jul-1989 #text\_change 18-Aug-2000  
 C:Accession: A32039  
 R:Horvath, A.; Kastin, A.J.  
 J. Biol. Chem. 264, 2175-2179, 1989  
 A:Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor  
 A:Reference number: A32039; MUID:89123285; PMID:2563371  
 A:Accession: A32039  
 A:Molecule type: protein  
 A:Residues: 1-4 <HOR>  
 A:Experimental source: brain  
 C:Superfamily: unassigned animal peptides  
 C:Keywords: amidated carboxyl end  
 F:4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 26.2%; Score 11; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4  
 DB 2 PL 3

# RESULT 12

I54357  
 schwannomin - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C:Accession: I54357  
 R:Huyh, D.P.; Nechiporuk, T.; Pulst, S.  
 Hum. Mol. Genet. 3, 1075-1079, 1994  
 A:Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are co-transcribed with the NF2 gene  
 A:Reference number: I54357; MUID:95072570; PMID:7981675  
 A:Accession: I54357  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <RES>  
 A:Cross-references: GB:L28838; NID:9454836; PIDN:AAA57150.1; PID:G601923  
 C:Genetics: 6/1  
 A:Gene: NF2

Query Match 26.2%; Score 11; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3  
 DB 1 VP 2

# RESULT 13

A60521  
 glycogen phosphorylase (BC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)  
 C:Species: Liza ramada  
 C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2003  
 C:Accession: A60521  
 R:Bonamusa, L.; Baanante, I.V.  
 Comp. Biochem. Physiol. B 95, 295-301, 1990  
 A:Title: Purification and characterization of glycogen phosphorylase B from skeletal muscle  
 A:Reference number: A60521; MUID:90227907; PMID:2109669  
 A:Accession: A60521  
 A:Molecule type: protein  
 A:Residues: 1-5 <BON>  
 C:Superfamily: glucan phosphorylase  
 C:Keywords: glycyltransferase; hexosyltransferase; phosphoprotein  
 F:3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experimental

Query Match 26.2%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3  
 DB 4 VP 5

# RESULT 14

I49421  
 laminin B1 - western wild mouse (fragment)  
 C:Species: Mus spretus (western wild mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C:Accession: I49421  
 R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maeraki, Y.; Nadeau, J.H.  
 Mamm. Genome 5, 349-355, 1994  
 A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
 A:Reference number: I48934; MUID:94319082; PMID:8043949  
 A:Accession: I49421  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A;Residues: 1-6 <RES>  
A;Cross-References: EMBL:U05736; NID:G497073; PIDN:AAB60477.1; PID:G642829

Query Match 26.2%; Score 11; DB 2; Length 6;  
Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 TSC 7  
Db 3 STC 5

RESULT 15  
E61491  
seed protein ws-5 - winged bean (fragment)  
C;Species: Psophocarpus tetragonolobus (winged bean)  
C;Date: 07-Oct-1994 #sssequence\_revision 07-Oct-1994 #text\_change 07-Oct-1994  
C;Accession: E61491  
R;Hirano, H.  
J. Protein Chem. 8, 115-130, 1989  
A;Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-dimensional gel electrophoresis  
A;Reference number: A61491; MUID:89351606; PMID:2765119  
A;Accession: E61491  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-7 <HIR>  
C;Keywords: glycoprotein; seed

Query Match 26.2%; Score 11; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3  
Db 3 VP 4

Search completed: September 5, 2004, 11:19:01  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:07:09 ; Search time 16 Seconds  
(without alignments)  
22.781 Million cell updates/sec

Title: US-09-761-636A-11

Perfect score: 42

Sequence: 1 CVPLTSC 7

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	13	31.0	6	1	EI01_LITRU	P82096 litoria rub
2	11	26.2	6	1	VP19_HSV1K	P23210 herpes simp
3	11	26.2	7	1	MNP1_LEPDE	P42984 leptinotars
4	11	26.2	7	1	TPFY_PACDA	P83455 pachymedusa
5	10	23.8	7	1	CARP_MYTED	P10420 mytilus edu
6	10	23.8	7	1	TV51_LITRU	P82065 litoria rub
7	9	21.4	5	1	EI03_LITRU	P82099 litoria rub
8	9	21.4	6	1	CIP1_MYTED	P13736 mytilus edu
9	9	21.4	6	1	CIP2_MYTED	P13737 mytilus edu
10	9	21.4	7	1	BRHP_CONIM	P58803 conus imper
11	8	19.0	5	1	BIOA_CITFR	P13071 citrobacter
12	8	19.0	5	1	PRCT_PERAM	P01373 periplaneta
13	8	19.0	6	1	TRPI_PSEPU	P36414 pseudomonas
14	8	19.0	7	1	CIA_ENTFA	P11932 enterococcu
15	8	19.0	7	1	UN06_PINES	P81675 pinus pinas
16	7	16.7	3	1	THYL_PIG	P01151 sus scrofa
17	7	16.7	4	1	DCML_PSECH	P19916 pseudomonas
18	7	16.7	4	1	RM01_YEAST	P36515 saccharomyc
19	7	16.7	4	1	TUFT_HUMAN	P01858 homo sapien
20	7	16.7	5	1	BP7_BOTIN	P30425 bothrops in
21	7	16.7	5	1	EI04_LITRU	P82100 litoria rub
22	7	16.7	5	1	PAP2_PARMA	P81864 pardachirus
23	7	16.7	5	1	SUGA_ACHDO	P19991 acheta dome
24	7	16.7	6	1	OMV_LEPDE	P42985 leptinotars
25	7	16.7	6	1	TMOF_SABRU	P41495 sarcophaga
26	7	16.7	7	1	ALL3_CARMA	P81806 carcinus ma
27	7	16.7	7	1	ALL4_CARMA	P81807 carcinus ma
28	7	16.7	7	1	ALL5_CARMA	P81808 carcinus ma
29	7	16.7	7	1	CCF1_ENTFA	P20104 enterococcu
30	7	16.7	7	1	CHOX_ALCSP	P16101 alcaligenes
31	7	16.7	7	1	E105_LITRU	P82101 litoria rub
32	7	16.7	7	1	FAR1_HEUT1	P41871 heliosoma tr
33	7	16.7	7	1	FAR4_PANRE	P41875 panagrellus

## ALIGNMENTS

RESULT 1  
EI01\_LITRU  
ID EI01\_LITRU STANDARD; PRT; 6 AA.  
AC P82096;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Electrin 1.  
OS Litoria rubella (Desert tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;  
OC Pelodyadinae; Litoria.  
OX NCBI\_TaxID=104895;  
RN [1]  
RP SEQUENCE.  
RC TISSUE-Skin secretion;  
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
RT "Peptides from the skin glands of the Australian buzzing tree frog  
RT Litori electrica. Comparison with the skin peptides from Litoria  
RT rubella.";  
RL Aust. J. Chem. 52:639-645 (1999).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
KW Amphibian defense peptide; Amidation.  
FT MOD RES 6  
SQ SEQUENCE 6 AA; 792 MW; 5693704772C9A000 CRC64;

Query Match 31.0%; Score 13; DB 1; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.4e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VPL 4  
Db 2 VPI 4

RESULT 2  
VP19\_HSV1K  
ID VP19\_HSV1K STANDARD; PRT; 6 AA.  
AC P23210;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Capsid assembly and DNA maturation protein (Virion protein UL38)  
DE (Capsid protein VP19C) (Fragment).  
GN UL38.  
OS Herpes simplex virus (type 1 / strain KOS).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91101287; PubMed=1846198;  
RA Flanagan W.M., Papavasiliou A.G., Rice M., Hecht L.B.,  
RA Silverstein S., Wagner E.K.;  
RT "Analysis of the herpes simplex virus type 1 promoter controlling the

RT expression of UL38, a true late gene involved in capsid assembly." ;  
 RL J. Virol. 65:769-786(1991).  
 CC -!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE  
 CC EMBEDDED. BINDS DNA.  
 CC -!- SIMILARITY: belongs to the herpesviruses VP19C family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; MS7646; AAA45830.1; -  
 KW Capsid assembly; Coat protein; DNA-binding.  
 FT NON TER 6  
 SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;  
 Query Match 26.2%; Score 11; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PL 4  
 Db 5 PL 6  
 RESULT 3  
 MNPL\_LPEDE STANDARD; PRT; 7 AA.  
 ID MNPL\_LPEDE  
 AC P42984;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Myotrophic neuromyosin 1 (Led-MNP-1).  
 OS Leptinotarsa decemlineata (Colorado potato beetle).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
 CC Phytophaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;  
 CC Chrysomelini; Leptinotarsa.  
 CC NCBI\_TaxID=7539;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RX MEDLINE=95380343; PubMed=7651886;  
 RA Spittaels K., Vankeerberghen A., Schoofs L., Torrekens S.,  
 RA Grauwels L., van Leuven F., de Loof A.;  
 RT "Identification, characterization, and immunological localization of  
 RT a novel myotrophic neuromyosin in the Colorado potato beetle,  
 RT Leptinotarsa decemlineata." ;  
 RL Peptides 16:365-374(1995).  
 CC -!- FUNCTION: Myotrophic peptide. Stimulates the contractions of the  
 CC oviduct.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 705 MW; 6D73768745B5DB0 CRC64;  
 Query Match 26.2%; Score 11; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PL 4  
 Db 5 PL 6  
 RESULT 4  
 TPFY\_PACDA STANDARD; PRT; 7 AA.  
 ID TPFY\_PACDA  
 AC P83455;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Tryptophyllin-1 (pdt-1).  
 OS Pachymedusa dactylosa (Giant Mexican leaf frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylinae;  
 CC Phyllomedusinae; Pachymedusa.  
 CC NCBI\_TaxID=75988;  
 RN [1]  
 RP SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF  
 RP PRO-7.  
 RC TISSUE=Skin secretion;  
 RC Chen T.B., Orr D.F., Shaw C.;  
 RA "Pachymedusa dactylosa tryptophyllin-1 (pdt-1): structural  
 RT characterization, pharmacological activity and cloning of precursor  
 RT cDNA." ;  
 RL Submitted (SEP-2002) to Swiss-Prot.  
 CC -!- FUNCTION: Myoactive. Has selective relaxing activity on vascular  
 CC smooth muscle.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Skin.  
 CC -!- MASS SPECTROMETRY: MW=809.2; METHOD=MALDI.  
 DR GO; GO:0005576; Extracellular; NAS.  
 DR GO; GO:0045986; P-negative regulation of smooth muscle contractility; NAS.  
 KW Amphibian defense peptide; Amidation; Hydroxylation.  
 FT MOD\_RES 3  
 FT MOD\_RES 3  
 FT MOD\_RES 7  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 794 MW; 7772D37DC776350 CRC64;  
 Query Match 26.2%; Score 11; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VP 3  
 Db 6 VP 7  
 RESULT 5  
 CARP\_MYTED STANDARD; PRT; 7 AA.  
 ID CARP\_MYTED  
 AC P10420;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-MAR-1989 (Rel. 10, Last annotation update)  
 DE Catch-relaxing peptide (CARP).  
 OS Mytilus edulis (Blue mussel).  
 CC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;  
 CC Mytiloidea; Mytilidae; Mytilus.  
 CC NCBI\_TaxID=6550;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=88052022; PubMed=3676797;  
 RA Hirata T., Kubota I., Takabatake I., Kawahara A., Shimamoto N.,  
 RA Muneoka Y.;  
 RT "Catch-relaxing peptide isolated from Mytilus pedal ganglia." ;  
 RL Brain Res. 422:374-376(1987).  
 CC -!- FUNCTION: This peptide exhibits both potentiating (contraction)  
 CC and inhibitory (relaxation) effects on the anterior byssus  
 CC retractor muscle.  
 DR PIR; A29342; ECMUCR.  
 KW Hormone; Amidation.  
 FT MOD\_RES 7  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 831 MW; 6734072687669DB0 CRC64;  
 Query Match 23.8%; Score 10; DB 1; Length 7;  
 Best Local Similarity 33.3%; Pred. No. 1.4e+05;  
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VPL 4  
 Db 2 MPM 4

```
RESULT 6
TY51_LITRU
ID TY51_LITRU STANDARD; PRT; 7 AA.
AC P82065;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophyllin 5.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella'. The skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: May act as a neuromodulator or neurotransmitter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC -!- MASS SPECTROMETRY: MW=965; METHOD=FAB.
KW Amphibian defense peptide; Amidation; Neuropeptide;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 23.8%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3
DB 2 IP 3

RESULT 7
EI03_LITRU
ID EI03_LITRU STANDARD; PRT; 5 AA.
AC P82099;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 3.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RA "Peptides from the skin glands of the Australian buzzing tree frog
RT Litori electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 21.4%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4

RESULT 8
CIPL_MYTED
ID CIPL_MYTED STANDARD; PRT; 6 AA.
AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide I (MIP I).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE=Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -!- FUNCTION: Inhibitory action on contractions in several molluscan
CC muscles.
CC -!- SIMILARITY: TO MIP II.
DR PIR; A27696; A27696.
KW Hormone; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 637 MW; 72C9C68775B81000 CRC64;

Query Match 21.4%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4

RESULT 9
CIP2_MYTED
ID CIP2_MYTED STANDARD; PRT; 6 AA.
AC P13737;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide II (MIP II).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE=Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -!- FUNCTION: Inhibitory action on contractions in several molluscan
CC muscles.
CC -!- SIMILARITY: TO MIP I.
DR PIR; B27696; B27696.
KW Hormone; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;

Query Match 21.4%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
```

```

Db          3 PM 4
RESULT 10
BRHP CONIM          STANDARD;          PRT;          7 AA.
AC P58603;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromheptapeptide Im.
OS Conus imperialis (Imperial cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=35631;
[1]
RN SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RP TISSUE=Venom; PubMed=9030520;
RX MEDLINE=97184108; Dykert J., Nielsen D.B., Gulyas J.,
RA Craig A.G., Jimenez E.C., Porter J., Rivier J.E., Olivera B.M.,
RA Abogadie F.C., Porter J., Rivier J.E., Olivera B.M.,
RA McIntosh J.M.;
RT "A novel post-translational modification involving bromination of
RT tryptophan. Identification of the residue, L-6-bromotryptophan, in
RT peptides from Conus imperialis and Conus radiatus venom.";
RL J. Biol. Chem. 272:4689-4698(1997).
CC -!- FUNCTION: Does not elicit gross behavioral symptoms when injected
CC centrally or peripherally in mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=853.19; METHOD=LSIMS.
DR PIR; A58512;
KW Bromination; Amidation; Pyrrolidone carboxylic acid.
FT DISULFID 2 7
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 BROMINATION.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 795 MW; 6EA37DC6D87EA6B0 CRC64;
Query Match 21.4%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 C 1
Db 2 C 2
RESULT 11
BIOA_CITFR
ID BIOA_CITFR          STANDARD;          PRT;          5 AA.
AC P13071;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
DE aminotransferase) (Fragment).
GN BIOA.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=8906280; PubMed=2971595;
RX Shuan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL Gene 67:203-211(1988).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-

```

```

CC diammonononanoate.
CC -!- COFACTOR: Pyridoxal phosphate.
CC -!- PATHWAY: Biotin biosynthesis.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21922; -; NOT_ANNOTATED_CDS.
DR PIR; I40697; I40697.
DR InterPro; IPR005814; Aminotrans_3.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW Biotin biosynthesis; Transferase; Aminotransferase;
KW Pyridoxal phosphate.
FT NON_TER 5
SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;
Query Match 19.0%; Score 8; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1.4e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 4 LTS 6
Db 1 MTT 3
RESULT 12
PRCT_PERAM
ID PRCT_PERAM          STANDARD;          PRT;          5 AA.
AC P01373;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Proctolin.
OS Periplaneta americana (American cockroach), and
OS Limulus polyphemus (Atlantic horseshoe crab), and
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OC NCBI_TaxID=6978; 6850; 6759;
RN [1]
RP SEQUENCE.
RC SPECIES=P.americana;
RX MEDLINE=76074708; PubMed=576;
RA Starratt A.N., Brown B.E.;
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT in insects.";
RL Life Sci. 17:1253-1256 (1975).
RN [2]
RP BIOLOGICAL SOURCE.
RC SPECIES=P.americana;
RX MEDLINE=8125865; PubMed=6113690;
RA O'Shea M., Adams M.E.;
RT "Pentapeptide (proctolin) associated with an identified neuron.";
RL Science 213:567-569(1981).
RN [3]
RP SEQUENCE.
RC SPECIES=L.polyphemus;
RX MEDLINE=90287800; PubMed=2356151;
RA Groome J.R., Tillinghast B.K., Townley M.A., Vetrovs A.,
RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA Shabanowitz J.;
RT "Identification of proctolin in the central nervous system of the
RT horseshoe crab, Limulus polyphemus.";
RL Peptides 11:205-211(1990).

```

```

RN RP SEQUENCE.
RC SPECIES=C.maenas; PubMed=2872661;
RX MEDLINE=86232789; PubMed=2872661;
RA Stangier J., Dirksen H., Keller R.;
RT "Identification and immunocytochemical localization of proctolin in
RL Pericardial organs of the shore crab, Carcinus maenas.";
RL Peptides 7:67-72(1986).
CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,
CC modulates visceral and skeletal muscle in many arthropods.
CC -!- TISSUE SPECIFICITY: Found in the lateral white neurons and in
CC the crab pericardial organs.
DR PIR; A01644; HOROHA.
DR PIR; A60411; A60411.
KW Neuropeptide.
SQ SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;

Query Match 19.0%; Score 8; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3
Db 3 LP 4

RESULT 13
TRPI_PSEPU STANDARD; PRT; 6 AA.
AC P36414.
DT 01-JUN-1994 (Rel. 29, Created)
DR 01-JUN-1994 (Rel. 29, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE HTH-type transcriptional regulator trpI (TrpEA operon transcriptional
DE activator) (Fragment).
GN TRPI.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PPG1 C15;
RX MEDLINE=89335826; PubMed=2503057;
RA Eberly L., Crawford I.P.;
RT "DNA sequence of the tryptophan synthase genes of Pseudomonas
RL putida.";
RL Biochimie 71:521-531(1989).
CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPA GENES ENCODING
CC THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE
CC INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE
CC TRPA OPERON, OVERLAPPING ITS OWN PROMOTER REGION.
CC -!- SIMILARITY: Contains 1 HTH lyseR-type DNA-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X13299; CAA31660.1; --
CC InterPro; IPR000847; HTH_LYSR.
CC PROSITE; PS50931; HTH_LYSR; PARTIAL.
CC KW Tryptophan biosynthesis; Transcription regulation; Activator;
CC DNA-binding.
FT NON TER 6
SQ SEQUENCE 6 AA; 683 MW; 77672AA1EDD6F000 CRC64;

Query Match 19.0%; Score 8; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VP 3
Db 5 LP 6

RESULT 14
CIA_ENTFA STANDARD; PRT; 7 AA.
ID CIA_ENTFA
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis) (CIA).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=87005252; PubMed=3093276;
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
RL CAM373.";
RL FEBS Lett. 206:69-72(1986).
CC -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC HARBORING PAM373.
CC -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR PIR; A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 19.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTS 6
Db 5 LAS 7

RESULT 15
UN06_PINPS STANDARD; PRT; 7 AA.
ID UN06_PINPS
AC P81675;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of needles (N141) (Fragment).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RX TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RL proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.6, its MW is: 25 kDa.
FT NON TER 1
FT NON TER 7
SQ SEQUENCE 7 AA; 823 MW; 69D76724486B5740 CRC64;

Query Match 19.0%; Score 8; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 VP 3  
: |  
Db 5 LP 6

Search completed: September 5, 2004, 11:17:30  
Job time : 17 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:13:09 ; Search time 57 Seconds  
(without alignments)  
38.748 Million cell updates/sec

Title: US-09-761-636A-11  
Perfect score: 42  
Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	17	40.5	7	12	Q66113	Q66113 cherry leaf
2	15	35.7	7	13	Q42564	Q42564 figu rubrip
3	12	28.6	5	13	P83308	P83308 gallus gall
4	12	28.6	7	12	Q67113	Q67113 influenzavi
5	11	26.2	7	4	Q8NH7	Q8nh7 homo sapien
6	11	26.2	7	8	Q8MFY6	Q8mfy6 taraxacum (
7	11	26.2	7	10	P93233	P93233 lycopersico
8	10	23.8	7	2	P70804	P70804 azotobacter
9	10	23.8	7	2	P72081	P72081 nocardia la
10	10	23.8	7	10	Q8GL12	Q8gl12 borrelia bu
11	10	23.8	7	10	Q9CSB3	Q9csb3 arabidopsis
12	10	23.8	7	15	Q07624	Q07624 rous sarcom
13	9	21.4	6	5	P83569	P83569 sepiia offic
14	9	21.4	7	2	O50556	O50556 actinobacil
15	9	21.4	7	2	Q54248	Q54248 streptomyce
16	9	21.4	7	11	O55184	O55184 rattus norv

17	9	21.4	7	12	Q9YVE3	Q9yve3 human adeno
18	9	21.4	7	12	Q9YIQ9	Q9yiq9 human adeno
19	9	21.4	7	12	Q9YIRO	Q9yiro human adeno
20	9	21.4	7	13	Q8JU20	Q8ju20 gallus gall
21	8	19.0	7	11	Q63668	Q63668 rattus norv
22	7	16.7	7	2	Q47029	Q47029 enterobacte
23	7	16.7	7	2	Q34028	Q34028 sphingomona
24	7	16.7	7	2	P83530	P83530 lactobacill
25	7	16.7	7	3	P83492	P83492 bionectria
26	7	16.7	7	5	Q721C0	Q721c0 caenorhabdi
27	7	16.7	7	8	P92214	P92214 amblyopyrum
28	7	16.7	7	8	P92393	P92393 hordeum vul
29	7	16.7	7	8	P92403	P92403 lophopyrum
30	7	16.7	7	8	P92427	P92427 peridictyon
31	7	16.7	7	8	P92430	P92430 aegilops ta
32	7	16.7	7	8	P92221	P92221 bromus iner
33	7	16.7	7	8	O98866	O98866 spinacia ol
34	7	16.7	7	8	P92425	P92425 pseudoroegn
35	7	16.7	7	8	P92381	P92381 hordeum bra
36	7	16.7	7	8	P92387	P92387 henrardia p
37	7	16.7	7	8	P92210	P92210 agropyron c
38	7	16.7	7	8	P92440	P92440 thinopyrum
39	7	16.7	7	8	P92218	P92218 australopyr
40	7	16.7	7	8	P92390	P92390 heteranthe
41	7	16.7	7	8	P92372	P92372 haynaldia v
42	7	16.7	7	8	P92442	P92442 taeniatheru
43	7	16.7	7	8	P92226	P92226 crithopsis
44	7	16.7	7	8	P92385	P92385 hordeum mar
45	7	16.7	7	8	P92421	P92421 psathyrosta

## ALIGNMENTS

RESULT 1  
Q66113 PRELIMINARY; PRT; 7 AA.  
AC Q66113; 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DS C-terminus of the viral replicase (Fragment).  
OS Cherry leaf roll virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;  
OC Nepovirus.  
OX NCBI\_TaxID=12615;  
RN [1] \_SEQUENCE FROM N.A.  
RP STRAIN=Walnut;  
RC STRAIN=Walnut;  
RA Borja M.;  
RL Thesis (1992), Biologia Molecular y Virologia Vegetal, CIT-INIA.  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Walnut;  
RX MEDLINE=96124520; PubMed=8560786;  
RA Borja M., Sanchez F., Rowhani A., Bruening G., Ponz F.;  
RT "Long, nearly identical untranslated sequences at the 3' terminal  
regions of the genomic RNAs of cherry leafroll virus (walnut  
strain)".  
RL Virus Genes 10:245-252(1995).  
DR EMBL: Z34265; CAA84019.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 800 MW; 7417672EBDC6D740 CRC64;

Query Match 40.5%; Score 17; DB 12; Length 7;  
Best Local Similarity 66.7%; Pred.No. 1e+06;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVP 3  
Db 4 CLP 6

```

RESULT 2
O42564 ID O42564 PRELIMINARY; PRT; 7 AA.
AC O42564;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN SCN8A.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97442476; PubMed=9295353;
RA Plummer N.W., McBurney M.W., Meisler M.H.;
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT two-domain protein in fetal brain and non-neuronal cells.";
RL J. Biol. Chem. 272:24008-24015(1997).
DR EMBL; U97673; AAB80916.1; -.
DR GO; GO:0005216; F-ion channel activity; IEA.
KW Ionic channel.
FT NON_TER 1
SQ SEQUENCE 7 AA; 730 MW; 75B72EA2C73772A0 CRC64;

Query Match 35.7%; Score 15; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPL 4
DB 1 VPL 3

RESULT 3
P83308 ID P83308 PRELIMINARY; PRT; 5 AA.
AC P83308;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FMRFamide-like neuropeptide (LPLRF-amide).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX TISSUE=Brain;
RX PubMed=6137771;
RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
RT "A novel active pentapeptide from chicken brain identified by
RT antibodies to FMRFamide.";
RL Nature 305:328-330(1983).
CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR GO; GO:0007218; P-neuropeptide signaling pathway; TAS.
KW Neuropeptide; Amidation.
FT MOD_RES 5
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 28.6%; Score 12; DB 13; Length 5;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPL 4
DB 1 LPL 3

RESULT 4
Q67113 ID Q67113 PRELIMINARY; PRT; 7 AA.
AC Q67113;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Influenza virus type A (Udorn/72) hemagglutinin (Seg 4) cDNA, 3' end
DE (Fragment).
OS Influenzavirus A.
OC Viruses; SSRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses.
OX NCBI_TaxID=197911;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81001892; PubMed=7407922;
RA Dhar R., Chanock R.M., Lai C.-J.;
RT "Nonviral oligonucleotides at the 5' terminus of cytoplasmic influenza
RT viral mRNA deduced from cloned complete genomic sequences.";
RL Cell 21:495-500(1980).
DR EMBL; M25045; AAA43202.1; -.
DR NON_TER 1
SQ SEQUENCE 7 AA; 834 MW; 605EB0544EA40030 CRC64;

Query Match 28.6%; Score 12; DB 12; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CV 2
DB 6 CI 7

RESULT 5
Q8NHH7 ID Q8NHH7 PRELIMINARY; PRT; 7 AA.
AC Q8NHH7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Mini-cistron.
GN NHE3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Malakooti J., Ramaswamy K.;
RT "Molecular cloning and characterization of the human Na+/H+ exchanger
RT NHE-3 gene promoter region.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282824; AAM53436.1; -.
SQ SEQUENCE 7 AA; 842 MW; 74072DC772D406F0 CRC64;

Query Match 26.2%; Score 11; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3
DB 3 VP 4

RESULT 6
Q8MFY6 ID Q8MFY6 PRELIMINARY; PRT; 7 AA.
AC Q8MFY6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

```

DE PsbA (Fragment).
GN PsbA.
OS Taraxacum (sect. Dioszegia) sp. 4310Hnew.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
OC Taraxacum.
OX NCBI_TaxID=154248;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Mes T.H.M.;
RT Reconstruction of the evolution of trnF pseudogenes."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY015477; AAC121591.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA; 675 MW; 687451B5A76DDB70 CRC64;

Query Match          26.2%; Score 11; DB 8; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PLTS 6
DB 3 PSTN 6

RESULT 7
P93233
ID P93233 PRELIMINARY; PRT; 7 AA.
AC P93233;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
DE (Fragment).
GN LE-ACS1B.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97351561; PubMed=9207843;
RA Oetiker J.H.; Olson D.C.; Shiu O.Y.; Yang S.F.;
RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate
RT synthase genes by elicitor in suspension cultures of tomato
RT (Lycopersicon esculentum).";
RL Plant Mol. Biol. 34:275-286(1997).
DR EMBL; U75692; AAC49682.1; -.
DR GO; GO:0016847; F:1-aminocyclopropane-1-carboxylate synthase . . .; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
KW Lyase.
FT NON_TER
SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match          26.2%; Score 11; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 3 PL 4

RESULT 8
P70804
ID P70804 PRELIMINARY; PRT; 7 AA.
AC P70804;

```

```

DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Algt protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A.; Ertesvag H.; Valla S.;
RT "A new Azotobacter vinelandii manuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -.
FT NON_TER
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A2D1AED0 CRC64;

Query Match          23.8%; Score 10; DB 2; Length 7;
Best Local Similarity 33.3%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CVPPTS 6
DB 1 CTVSSS 6

RESULT 9
P72081
ID P72081 PRELIMINARY; PRT; 7 AA.
AC P72081;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 3'-methylcephem hydroxylase (Fragment).
GN CEFF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96009872; PubMed=7557411;
RA Coque J.; Perez-Llarena F.J.; Enguita F.J.; Fuente J.L.; Martin J.F.;
RA Litás P.;
RT "Characterization of the cmcH genes of Nocardia lactamdurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL; Z21682; CAA79797.1; -.
FT NON_TER
SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match          23.8%; Score 10; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LTS 6
DB 5 VTS 7

RESULT 10
Q8GL12
ID Q8GL12 PRELIMINARY; PRT; 7 AA.
AC Q8GL12;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

```

DE PF-50 protein (Fragment).  
GN PF-50.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OG Plasmid group cp32-9.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N40;  
RA Stevenson B., Miller J.C.;  
RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32  
prophages: conservation amidst diversity.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY142100; AA017911.1; -;  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid.  
FT NON\_TER  
SQ SEQUENCE 1 1 849 MW; 6337244330569ED0 CRC64;  
Query Match 23.8%; Score 10; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CV 2  
Db 1 CM 2  
RESULT 11  
O9C5B3 PRELIMINARY; PRT; 7 AA.  
ID Q9C5B3  
AC Q9C5B3  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN DIDI 10A-2B.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Root;  
RX MEDLINE=21171025; PubMed=11277426;  
RA Vercauteren I., Van Der Schueren E., Van Montagu M., Gheysen G.;  
RT "Arabidopsis thaliana genes expressed in the early compatible  
interaction with root-knot nematodes.";  
RL Mol. Plant Microbe Interact. 14:288-299 (2001).  
DR EMBL; AJ286350; CAB71014.2; -;  
FT NON\_TER  
SQ SEQUENCE 1 1 719 MW; 6732C7287EB325D0 CRC64;  
Query Match 23.8%; Score 10; DB 10; Length 7;  
Best Local Similarity 33.3%; Pred. No. 1e+06; Indels 1; Gaps 0;  
Matches 1; Conservative 1; Mismatches 1; Indels 1; Gaps 0;  
QY 5 TSC 7  
Db 1 SKC 3  
RESULT 12  
Q07624 PRELIMINARY; PRT; 7 AA.  
ID Q07624  
AC Q07624;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE UORF1.

OS Rous sarcoma virus (strain Prague C).  
OC Viruses; Retroviral viruses; Retroviridae; Alpharetrovirus.  
OX NCBI\_TaxID=11888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93010967; PubMed=1327749;  
RA Donze O., Spahr P.F.;  
RT "Role of the open reading frames of Rous sarcoma virus leader RNA in  
translation and genome packaging.";  
RL EMBO J. 11:3747-3757 (1992).  
DR EMBL; X67587; CA47862.1; -;  
SQ SEQUENCE 7 AA; 672 MW; 776045A7687DD6F0 CRC64;  
Query Match 23.8%; Score 10; DB 15; Length 7;  
Best Local Similarity 50.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VP 3  
Db 6 IP 7  
RESULT 13  
P83569 PRELIMINARY; PRT; 6 AA.  
ID P83569  
AC P83569;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Sperm attracting peptide SepsAP.  
OS Sepia officinalis (Common cuttlefish).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
OC Decapodiformes; Sepioidea; Sepiidae; Sepia.  
OX NCBI\_TaxID=6610;  
RN [1]  
RP SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND  
AMIDATION.  
RC TISSUE=Egg;  
RX PubMed=12207899;  
RA Zatylny C., Marvin L., Gagnon J., Henry J.;  
RT "Fertilization in Sepia officinalis: the first mollusk sperm-  
attracting peptide.";  
RL Biochem. Biophys. Res. Commun. 296:1186-1193 (2002).  
CC -!- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE  
COLLISION.  
CC -!- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES FURING  
VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED  
DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.  
CC ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.  
CC -!- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.  
KW Amidation.  
FT MOD\_RES  
SQ SEQUENCE 6 AA; 597 MW; 72C8676AA0470000 CRC64;  
Query Match 21.4%; Score 9; DB 5; Length 6;  
Best Local Similarity 50.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 PL 4  
Db 1 PI 2  
RESULT 14  
O50556 PRELIMINARY; PRT; 7 AA.  
ID O50556  
AC O50556;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GLYA (Fragment).  
GN GLYA.  
OS Actinobacillus actinomycetemcomitans (Haemophilus

```

OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384;
RX MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans.";
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 21.4%; Score 9; DB 2; Length 7;
Best Local Similarity 33.3%; Pred. No. 1e+06;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VPL 4
Db 3 LPV 5

RESULT 15
Q54248
ID Q54248 PRELIMINARY; PRT; 7 AA.
AC Q54248;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RplO protein (Fragment).
GN RPL0.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RX MEDLINE=20011291; PubMed=10542330;
RA Poehling S., Piepersberg W., Wehmeier U.F.;
RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL Biochim. Biophys. Acta 1447:298-302(1999).
DR EMBL; X95915; CAA65160.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 21.4%; Score 9; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VPLT 5
Db 1 VTVT 4

```

Search completed: September 5, 2004, 11:18:34  
 Job time : 58 secs

**his Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:06:29 ; Search time 63 Seconds  
(without alignments)  
31.394 Million cell updates/sec

Title: US-09-761-636A-11

Perfect score: 42

Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 92273

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	42	100.0	7	AAU04530	AAU04530 VEGF base
2	31	73.8	7	AAW13421	AAW13421 Kidney ho
3	31	73.8	7	AAAB12007	AAAB12007 Kidney ho
4	31	73.8	7	AAE11813	AAE11813 Phage pep
5	31	73.8	7	AAU10724	AAU10724 Kidney ho
6	31	73.8	7	ABU59533	ABU59533 Kidney re
7	31	73.8	7	ADC33700	ADC33700 Kidney ce
8	31	73.8	7	ADC33699	ADC33699 Kidney ce
9	29	69.0	6	AAU04531	AAU04531 VEGF base
10	28	66.7	7	ABJ00550	ABJ00550 B lymphoc
11	28	66.7	7	ABG33862	ABG33862 B lymphoc
12	26	61.9	7	AAE11489	AAE11489 Cadherin-
13	25	59.5	7	AAE11489	AAE11489 PB-cadher
14	25	59.5	7	AAE11489	AAE11489 PB-cadher
15	25	59.5	7	AAE11489	AAE11489 PB-cadher
16	24	57.1	7	AAE11489	AAE11489 PB-cadher
17	24	57.1	7	AAE11489	AAE11489 PB-cadher
18	23	54.8	6	AAE11449	AAE11449 IGF-1 ana
19	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
20	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
21	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
22	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
23	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
24	23	54.8	7	AAE11449	AAE11449 IGF-1 ana
25	23	54.8	7	AAE11449	AAE11449 IGF-1 ana

26	23	54.8	7	5	ABB47076	Abb47076 Desmocoll
27	22	52.4	4	2	AAR15772	Aar15772 Farnesyl-
28	22	52.4	4	2	AAR49769	Aar49769 Farnesyl-
29	22	52.4	4	2	AAR77833	Aar77833 Farnesyl-
30	22	52.4	4	2	AAW04459	Aaw04459 Farnesyl
31	22	52.4	5	2	AAW67428	Aaw67428 HCV pepti
32	22	52.4	6	6	ABJ37236	Abj37236 Rhodopsin
33	22	52.4	7	2	AAR36871	Aar36871 Insulin-l
34	22	52.4	7	2	AAR43615	Aar43615 Peptide d
35	22	52.4	7	2	AAW67430	Aaw67430 HCV pepti
36	22	52.4	7	3	AAV64307	Aay64307 Cadherin-
37	22	52.4	7	3	AAV64292	Aay64292 Cadherin-
38	22	52.4	7	3	AAV62135	Aay62135 Cadherin-
39	22	52.4	7	3	AAV63266	Aay63266 Protocadh
40	22	52.4	7	5	ABJ04536	Abj04536 Molt-4 le
41	22	52.4	7	5	ABJ04536	Abj04536 Molt-4 le
42	21	50.0	7	2	AAW60386	Aaw60386 Tumour ho
43	21	50.0	7	2	AAW93712	Aaw93712 Human bre
44	21	50.0	7	3	AAW03612	Aaw03612 Aspergill
45	21	50.0	7	3	AAV64262	Aay64262 Cadherin-

## ALIGNMENTS

RESULT 1  
AAU04530  
ID AAU04530 standard; peptide; 7 AA.  
XX  
AC AAU04530;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE VEGF based monocyclic peptide 8.  
XX  
KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
KW diabetic retinopathy; chronic inflammation; cyclic.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..7 /note= "This bond cyclises the peptide"  
XX  
XX WO200152875-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 18-JAN-2001; 2001WO-US001533.  
XX  
PR 18-JAN-2001; 2000US-0176293P.  
XX  
PR 16-MAY-2000; 2000US-0204590P.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Achen MG, Hughes RA, Stacker S, Cendron A;  
XX  
XX WPI; 2001-442248/47.  
XX  
PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
PT residues.  
XX  
PS Claim 49; Page 32; 102pp; English.  
XX  
CC The sequence represents a monomeric monocyclic peptide of the invention,  
CC whose 3-dimensional structure is modelled on the expose loop of human  
CC VEGFD (vascular endothelial growth factor). The invention relates to a  
CC method of producing a monomeric monocyclic peptide by a measuring beta-  
CC beta carbon separation distances on opposite antiparallel strands of a

CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 42; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTSC 7  
 DB 1 CVPLTSC 7

RESULT 2  
 AAW13421  
 ID AAW13421 standard; peptide; 7 AA.

AC AAW13421;  
 XX  
 XX  
 DT 15-JAN-1998 (first entry)  
 XX  
 DE Kidney homing peptide.

XX Kidney homing peptide; in vivo panning; screening; phage display;  
 KW drug delivery.

XX Synthetic.

OS  
 XX W09710507-A1.

PN  
 XX 20-MAR-1997.

XX 10-SEP-1996; 96WO-US014600.

XX 11-SEP-1995; 95US-00526708.

PR 11-SEP-1995; 95US-00526710.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX  
 PI Ruoslahti E, Pasqualini R;

XX WPI; 1997-202359/18.

XX Obtaining compound that homes to selected organ or tissue - by in vivo  
 PT panning method, specifically to identify brain, kidney, angiogenic  
 PT vasculature or tumour tissue homing peptide(s).

XX Claim 16; Page 68; 75pp; English.

XX This synthetic peptide is a claimed example of a kidney-homing peptide  
 CC that was identified using a novel method for obtaining molecules that  
 CC home to a selected organ or tissue. This in vivo panning method typically

CC involves administering a phage display library to a subject, and  
 CC identifying expressed peptides which home to the desired organ or tissue,  
 CC e.g. brain, kidney, angiogenic vascular tissue or tumour tissue. The  
 CC isolated peptides (see AAW13412-52, AAW1181-86) can be used to target  
 CC e.g. drugs, toxins or labels to the selected organ/tissue (claimed) or to  
 CC identify and/or isolate target molecules (claimed). The peptides can be  
 CC directly identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo

XX Sequence 7 AA;

Query Match 73.8%; Score 31; DB 2; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7  
 DB 1 CLPVASC 7

RESULT 3

AAB12007  
 ID AAB12007 standard; peptide; 7 AA.

XX  
 AC AAB12007;

XX 17-OCT-2000 (first entry)

XX Kidney homing peptide # 1.

XX Kidney; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX Mus sp.

XX Key Location/Qualifiers

FT Disulfide-bond 1..7 /note= "Can optionally form a cyclic peptide"

XX US6068829-A.

XX 30-MAY-2000.

XX 23-JUN-1997; 97US-00862855.

XX 11-SEP-1995; 95US-00526710.

PR 10-MAR-1997; 97US-00813273.

XX (BURN-) BURNHAM INST.

XX Pasqualini R, Ruoslahti E;

XX WPI; 2000-410850/35.

XX Identifying and recovering organ homing molecules or peptides by in vivo  
 PT panning comprises administering a library of diverse peptides linked to a  
 PT tag which facilitates recovery of these peptides.

XX Example 2; Col 18; 20pp; English.

XX The present sequence is a mouse kidney homing peptide. This sequence was  
 CC identified by using in vivo panning to screen a library of potential  
 CC organ homing molecules. The present sequence can be used to direct a  
 CC moiety to a kidney tissue, by linking the moiety to the present sequence.  
 CC Examples of potential moieties are drugs, toxins or a detectable label

XX Sequence 7 AA;

Query Match 73.8%; Score 31; DB 3; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7



```

Db      |.:|:| |
        1 CLPVASC 7

RESULT 4
AAE11813
ID  AAE11813 standard; peptide; 7 AA.
XX
XX  AAE11813;
AC
XX
XX  11-SEP-2003 (revised)
DT  18-DEC-2001 (first entry)
XX
XX  Phage peptide #21 targetted to kidney.
DE
XX
XX  Enriched library fraction; brain; kidney; tumour; panning; diagnostic;
KW  molecular medicine; drug delivery; peptidomimetic; pharmaceutical.
XX
XX  unidentified bacteriophage.
OS
OS  Unidentified.
XX
XX  US6296832-B1.
PN
XX
XX  02-OCT-2001.
PD
XX
XX  08-JAN-1999; 99US-00226985.
PF
XX
XX  11-SEP-1995; 95US-00526710.
PR
XX  10-MAR-1997; 97US-00813273.
PR
XX  23-JUN-1997; 97US-00862855.
XX
XX  (BURN-) BURNHAM INST.
PA
XX
XX  Ruoslahti E, Pasqualini R;
PI
XX
XX  WPI; 2001-610691/70.
DR
XX
XX  Enriched library fraction comprising molecules recovered by in vivo
PT  panning that selectively home to a selected organ or tissue useful for
PT  treating disease or in diagnostic methods.
XX
XX  Example 2; Col 18; 21pp; English.
XX
XX  The invention relates to an enriched library fraction containing
CC  molecules that selectively home to a selected organ or tissue such as
CC  brain, kidney or tumour recovered by in vivo panning. The invention
CC  generally relates to the field of molecular medicine, drug delivery and
CC  to a method of in vivo panning for identifying a molecule that homes to a
CC  specific organ. The molecules, e.g., peptides, peptidomimetics, proteins
CC  and fragments of proteins contained in an enriched library fraction may
CC  be administered to a subject as part of a pharmaceutical composition to
CC  treat disease or in diagnostic methods. The present sequence is a peptide
CC  from bacteriophage targetted to kidney. (Updated on 11-SEP-2003 to
CC  standardise OS field)
XX
XX  Sequence 7 AA;
SQ

Query Match      73.8%; Score 31; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy  1 CVPLTSC 7
    |.:|:| |
Db  1 CLPVASC 7

RESULT 5
AAU10724
ID  AAU10724 standard; peptide; 7 AA.
XX
XX  AAU10724;
AC
XX
XX  12-MAR-2002 (first entry)
DT

```

```

XX
DE
XX
KW  Organ targeting; tissue targeting; cancer; tumour homing molecule;
KW  delivery of target molecule; kidney homing peptide.
XX
OS  Synthetic.
XX
XX  US6306365-B1.
PN
XX
XX  23-OCT-2001.
PD
XX
XX  08-JAN-1999; 99US-00227906.
PF
XX
XX  11-SEP-1995; 95US-00526710.
PR
XX  10-MAR-1997; 97US-00813273.
PR
XX  23-JUN-1997; 97US-00862855.
XX
XX  (BURN-) BURNHAM INST.
PA
XX
XX  Ruoslahti E, Pasqualini R;
PI
XX
XX  WPI; 2002-040196/05.
DR
XX
XX  Recovering molecules that home to an organ or tissue, useful for
PT  identifying molecules that home to a specific organ or tissue, e.g.
PT  identifying a tumor homing molecule to identify the presence of cancer,
PT  by in vivo panning of a library.
XX
XX  Example 2; Col 18; 21pp; English.
XX
XX  The present invention relates to a method of recovering molecules that
CC  home to a selected organ or tissue. The method comprises administering to
CC  the subject the library of diverse molecules, collecting a sample of the
CC  selected organ or tissue (e.g. brain or kidney), and recovering from the
CC  sample several molecules that home to the selected organ or tissue. The
CC  method is useful for identifying molecules, particularly useful for
CC  screening large number of molecules (e.g. peptides), that home to a
CC  specific organ. The identified molecule is useful for e.g. raising an
CC  antibody specific for a target molecule, targeting a desired moiety (e.g.
CC  drug, toxin or detectable label) to the selected organ. Specifically, the
CC  method is useful for identifying the presence of cancer in a subject by
CC  linking an appropriate moiety to a tumour homing molecule. The present
CC  method provides a direct means for identifying molecules that
CC  specifically home to a selected organ and, therefore provides a
CC  significant advantage over previous methods, which require that a
CC  molecule identified using an in vitro screening method subsequently be
CC  examined to determine if it maintains its specificity in vivo. AAU10724-
CC  AAU10738 represent kidney homing peptides described in the present
CC  invention
XX
XX  Sequence 7 AA;
SQ

Query Match      73.8%; Score 31; DB 5; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy  1 CVPLTSC 7
    |.:|:| |
Db  1 CLPVASC 7

RESULT 6
ABU59533
ID  ABU59533 standard; peptide; 7 AA.
XX
XX  ABU59533;
AC
XX
XX  22-APR-2003 (first entry)
DT
XX
XX  Kidney receptor targeting peptide #1.
XX
XX  Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;
KW

```

KW cathepsin-D substrate; peptides; neuroreceptor; adrenal receptor;  
 KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;  
 KW tumour; cationic cancer-targeting peptide.  
 XX Synthetic.  
 OS  
 XX US2002041898-A1.  
 PN  
 XX 11-APR-2002.  
 PD  
 XX 25-JUL-2001; 2001US-00912609.  
 PF  
 XX 05-JAN-2000; 2000US-00478124.  
 PR 31-OCT-2000; 2000US-00703474.  
 XX  
 XX (UNGE/) UNGER E C.  
 PA (MATS/) MATSUNAGA T O.  
 PA (RAMA/) RAMASWAMI V.  
 PA (RONA/) ROMANOWSKI M J.  
 XX  
 XX Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;  
 PI WPI; 2003-208921/20.  
 XX  
 DR Targeted delivery system comprising a bioactive agent homogeneously  
 XX dispersed in a targeted matrix is especially useful in cancer therapy.  
 PT  
 PT Claim 25; Page 38; 46pp; English.  
 XX  
 XX The invention relates to a composition comprising a bioactive agent  
 CC homogeneously dispersed in a targeted matrix (polymer and targeting  
 CC ligand). Also included are a targeted matrix for use as a delivery  
 CC vehicle comprising a polymer associated with a targeting ligand,  
 CC enhancing the bioavailability of an agent comprising administration of the  
 CC the composition and treating cancer comprising administration of the  
 CC novel composition. The method is useful for targeted delivery of a drug,  
 CC especially in cancer therapy. The targeting ligand may be a peptide.  
 CC Examples of targeting peptides are disclosed including cathepsin-D  
 CC substrate peptides, peptides targeting receptors in the brain and kidney,  
 CC peptides recognising fibronectin- and vitronectin-binding integrins,  
 CC peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g., antibodies,  
 CC peptides targeting the angiogenic endothelium of solid tumours, tissue  
 CC specific peptides (e.g. of lung, skin, pancreas, intestine, uterus,  
 CC adrenal gland and retina), and cationic cancer-targeting peptides. The  
 CC present sequence is a peptide targeting ligand disclosed in the invention  
 XX  
 XX SQ Sequence 7 AA;  
 Query Match 73.8%; Score 31; DB 6; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 Db 1 CLPVASC 7  
 |.:|:|  
 RESULT 7  
 ADC33700  
 ID ADC33700 standard; peptide; 7 AA.  
 XX  
 XX ADC33700;  
 AC  
 XX 18-DEC-2003 (first entry)  
 DT  
 XX Kidney cell targeted peptide SEQ ID NO:4.  
 DE  
 XX chimeric retrovirus envelope protein; ecotropic envelope protein;  
 KW cytosstatic; gene therapy; cancer.  
 XX Synthetic.  
 OS  
 XX WO2003076596-A2.  
 PN

XX 18-SEP-2003.  
 PD  
 XX 07-MAR-2003; 2003WO-US007323.  
 PF  
 XX 08-MAR-2002; 2002US-0362655P.  
 PR  
 XX (UYMA-) UNIV MASSACHUSETTS.  
 PA  
 XX Green MR, Gollan TJ;  
 PI WPI; 2003-722332/68.  
 XX  
 DR New chimeric retrovirus envelope protein comprising an ecotropic envelope  
 XX protein and a heterologous short peptide ligand inserted within the  
 PT ecotropic envelope protein useful for treating cancer.  
 PT  
 XX Disclosure; SEQ ID NO 4; 42pp; English.  
 PS  
 XX The present invention describes a chimeric retrovirus envelope protein  
 CC comprising an ecotropic envelope protein and a heterologous short  
 CC peptide ligand inserted within the ecotropic envelope protein. Also  
 CC described: (1) a nucleic acid molecule comprising a sequence encoding the  
 CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic  
 CC acid sequence encoding the chimeric envelope protein; (3) a recombinant  
 CC retroviral particle comprising a chimeric envelope protein comprising a  
 CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)  
 CC identifying a nucleic acid sequence encoding the chimeric envelope  
 CC protein that alters viral tropism; (5) delivering a nucleic acid sequence  
 CC to a cell; and (6) treating cancer. (1) has cytostatic activity and can  
 CC be used in gene therapy. The chimeric retrovirus envelope protein is  
 CC useful for treating cancer, which comprises providing a cancer cell, e.g.  
 CC human cancer cell and infecting the cancer cell with a virus, e.g.  
 CC retrovirus comprising the chimeric envelope protein comprising a  
 CC heterologous short peptide ligand and a therapeutically useful gene, e.g.  
 CC encoding thymidine kinase. The present sequence represents a kidney cell  
 CC targeted peptide, which is given in the exemplification of the present  
 CC invention.  
 XX  
 XX SQ Sequence 7 AA;  
 Query Match 73.8%; Score 31; DB 7; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 Db 1 CLPVASC 7  
 |.:|:|  
 RESULT 8  
 ADC33699  
 ID ADC33699 standard; peptide; 7 AA.  
 XX  
 XX ADC33699;  
 AC  
 XX 18-DEC-2003 (first entry)  
 DT  
 XX Kidney cell targeted peptide SEQ ID NO:3.  
 DE  
 XX chimeric retrovirus envelope protein; ecotropic envelope protein;  
 KW cytosstatic; gene therapy; cancer.  
 XX Synthetic.  
 OS  
 XX WO2003076596-A2.  
 PN  
 XX 18-SEP-2003.  
 PD  
 XX 07-MAR-2003; 2003WO-US007323.  
 PF  
 XX 08-MAR-2002; 2002US-0362655P.  
 PR  
 XX

PA (UYMA-) UNIV MASSACHUSETTS.  
 XX Green MR, Gollan TJ;  
 PI WPI; 2003-722332/68.  
 DR  
 XX  
 XX  
 PT New chimeric retrovirus envelope protein comprising an ecotropic envelope  
 PT protein and a heterologous short peptide ligand inserted within the  
 PT ecotropic envelope protein useful for treating cancer.  
 XX  
 PS Disclosure; SEQ ID NO 3; 42pp; English.  
 XX  
 CC The present invention describes a chimeric retrovirus envelope protein  
 CC (1) comprising an ecotropic envelope protein and a heterologous short  
 CC peptide ligand inserted within the ecotropic envelope protein. Also  
 CC described: (1) a nucleic acid molecule comprising a sequence encoding the  
 CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic  
 CC acid sequence encoding the chimeric envelope protein; (3) a recombinant  
 CC retroviral particle comprising a chimeric envelope protein comprising a  
 CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)  
 CC identifying a nucleic acid sequence encoding the chimeric envelope  
 CC protein that alters viral tropism; (5) delivering a nucleic acid sequence  
 CC to a cell; and (6) treating cancer. (1) has cytostatic activity and can  
 CC be used in gene therapy. The chimeric retrovirus envelope protein is  
 CC useful for treating cancer, which comprises providing a cancer cell, e.g.  
 CC human cancer cell and infecting the cancer cell with a virus, e.g.  
 CC retrovirus comprising the chimeric envelope protein comprising a  
 CC heterologous short peptide ligand and a therapeutically useful gene, e.g.  
 CC encoding thymidine kinase. The present sequence represents a kidney cell  
 CC targeted peptide, which is given in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 7 AA;  
 Query Match 73.8%; Score 31; DB 7; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 Db 1 CLPVASC 7  
 RESULT 9  
 AAU04531  
 ID AAU04531 standard; peptide; 6 AA.  
 AC  
 AC AAU04531;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE VEGF based monocyclic peptide 9.  
 DE  
 DE Human; VEGF; vascular endothelial growth factor; angiogenesis;  
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
 KW diabetic retinopathy; chronic inflammation; cyclic.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..6  
 FT /note= "This bond cyclises the peptide"  
 FT  
 XX  
 XX WO200152875-A1.  
 XX  
 XX 26-JUL-2001.  
 XX  
 XX 18-JAN-2001; 2001WO-US001533.  
 XX  
 XX 18-JAN-2000; 2000US-0176293P.  
 XX  
 XX 16-MAY-2000; 2000US-0204590P.  
 XX

PA (LUDW-) LUDWIG INST CANCER RES.  
 XX Achen MG, Hughes RA, Stacker S, Cendron A;  
 PI WPI; 2001-442248/47.  
 XX  
 XX  
 PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
 PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
 PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
 PT residues.  
 XX  
 PS Claim 49; Page 32; 102pp; English.  
 XX  
 CC The sequence represents a monomeric monocyclic peptide of the invention,  
 CC whose 3-dimensional structure is modelled on the exposed loop of human  
 CC VEGF (vascular endothelial growth factor). The invention relates to a  
 CC method of producing a monomeric monocyclic peptide by a measuring beta-  
 CC beta carbon separation distances on opposite antiparallel strands of a  
 CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis, in a mammal with a condition  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 6 AA;  
 Query Match 69.0%; Score 29; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CVPLT 5  
 Db 1 CVPLT 5  
 RESULT 10  
 ABJ00550  
 ID ABJ00550 standard; peptide; 7 AA.  
 XX  
 AC ABJ00550;  
 XX  
 DT 05-SEP-2002 (first entry)  
 XX  
 DE B lymphocyte stimulator protein binding peptide #1.  
 XX  
 KW B lymphocyte stimulator protein binding protein; BLYS; immune disease;  
 KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
 KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
 KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
 KW immunosuppressive; nephrotropic; antirheumatic; antiarthritic;  
 KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;  
 KW antiasthmatic; antiallergic; thyromimetic; antinaemic; haemostatic;  
 KW dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;  
 KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
 XX  
 OS Unidentified.

XX FH Key Location/Qualifiers  
 XX FT Misc-difference 2  
 XX FT /label= Phe, Trp, Tyr  
 XX FT Misc-difference 4  
 XX FT /label= Pro, Tyr  
 XX  
 XX WO200216411-A2.  
 XX  
 XX PD 28-FEB-2002.  
 XX  
 XX PF 17-AUG-2001; 2001WO-US025850.  
 XX  
 XX PR 18-AUG-2000; 2000US-0226700P.  
 XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX PI Beltzer JP, Potter DM, Fleming TL, Rosen CA;  
 XX  
 XX DR WPI; 2002-499775/53.  
 XX  
 XX The treatment of various diseases e.g. rheumatoid arthritis, comprises  
 XX administering B Lymphocyte stimulator binding polypeptide.  
 XX  
 XX PS Claim 69; Page 233; 387pp; English.  
 XX  
 XX The present invention relates to the treatment, prevention or  
 XX amelioration of a disease or disorder associated with: aberrant B  
 XX lymphocyte stimulator (Blys), Blys receptor expression or activity; cells  
 XX of haematopoietic origin; or proliferative disease; and reducing,  
 XX inhibiting or stimulating immunoglobulin production, B cell proliferation  
 XX and graft rejection involving administration of Blys binding polypeptide.  
 XX The Blys binding polypeptides are used in the treatment, prevention or  
 XX amelioration of diseases such as immune system diseases, proliferative  
 XX diseases, diseases of cells of hematopoietic origin, graft rejection,  
 XX allergies, infectious diseases, arteriosclerosis, inflammatory disorders,  
 XX hyperagglutination, blood clotting disorders, ischaemia, and  
 XX neurodegenerative diseases. The present sequence is a B lymphocyte  
 XX stimulator protein binding peptide  
 XX  
 XX SQ Sequence 7 AA;  
 XX  
 XX Query Match 66.7%; Score 28; DB 5; Length 7;  
 XX Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 XX  
 XX QY 1 CVPLTSC 7  
 XX Db 1 CXPXTGC 7  
 XX  
 XX RESULT 11  
 XX ABG33862  
 XX ID ABG33862 standard; peptide; 7 AA.  
 XX  
 XX AC ABG33862;  
 XX  
 XX DT 15-JUL-2002 (first entry)  
 XX  
 XX DE B Lymphocyte Stimulator (Blys) binding peptide #436.  
 XX  
 XX KW B Lymphocyte Stimulator protein; B Lymphocyte Stimulator binding peptide;  
 XX Blys; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;  
 XX synovial fluid; saliva; mucus.  
 XX  
 XX OS Synthetic.  
 XX  
 XX XX WO200216412-A2.  
 XX  
 XX PD 28-FEB-2002.  
 XX  
 XX PF 17-AUG-2001; 2001WO-US025891.  
 XX

PR 18-AUG-2000; 2000US-0226489P.  
 XX  
 XX PA (DYAX-) DYAX CORP.  
 XX  
 XX PI Beltzer JP, Potter MD, Fleming TJ, Ladner RC;  
 XX  
 XX DR WPI; 2002-351647/38.  
 XX  
 XX New B-lymphocyte stimulator binding polypeptide useful in detecting or  
 XX isolating Blys or Blys-like polypeptide comprises a specified amino acid  
 XX sequence.  
 XX  
 XX PS Disclosure; Page 132; 269pp; English.  
 XX  
 XX The invention relates to a B Lymphocyte Stimulator (Blys) binding  
 XX polypeptide. Blys binding peptides bind Blys or Blys-like proteins  
 XX reversibly or irreversibly. The binding peptides are used in detection,  
 XX isolation and/or purification of Blys in a solution such as water or a  
 XX buffer solution, as well as any fluid and/or cell obtained from an  
 XX individual biological fluid, body tissue, body cell, cell line, tissue  
 XX culture or other source containing Blys or Blys-like polypeptides. The  
 XX biological fluids include sera, plasma, lymph, blood, blood fraction,  
 XX urine, synovial fluid, spinal fluid, saliva and mucous. Sequences  
 XX ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and  
 XX ABG33852-33862 represent Blys binding peptides of the invention  
 XX  
 XX SQ Sequence 7 AA;  
 XX  
 XX Query Match 66.7%; Score 28; DB 5; Length 7;  
 XX Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 XX  
 XX QY 1 CVPLTSC 7  
 XX Db 1 CXPXTGC 7  
 XX  
 XX RESULT 12  
 XX AAY61489  
 XX ID AAY61489 standard; peptide; 7 AA.  
 XX  
 XX AC AAY61489;  
 XX  
 XX DT 02-MAR-2000 (first entry)  
 XX  
 XX DE Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1375.  
 XX  
 XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 XX inhibition; cadherin extracellular domain; cell adhesion recognition;  
 XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 XX cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 XX cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 XX desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 XX rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 XX neurological disease; cyclic.  
 XX  
 XX OS Synthetic.  
 XX  
 XX XX Homo sapiens.  
 XX  
 XX FH Key Location/Qualifiers  
 XX FT Disulfide-bond 1. .7  
 XX  
 XX PN WO9957149-A2.  
 XX  
 XX XX 11-NOV-1999.  
 XX  
 XX XX 05-MAY-1999; 99WO-CA000363.  
 XX  
 XX XX 05-MAY-1998; 98US-00073040.  
 XX  
 XX PR 06-NOV-1998; 98US-00187859.  
 XX  
 XX PR 20-JAN-1999; 99US-00234395.  
 XX  
 XX PR 08-MAR-1999; 99US-00264516.  
 XX

PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX Claim 36; Page 172; 252pp; English.  
 XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in  
 CC the exemplification of the present invention  
 XX Sequence 7 AA;  
 SQ  
 Query Match 61.9%; Score 26; DB 3; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 Db 1 CEPKTGC 7  
 RESULT 13  
 AAY62764  
 ID AAY62764 standard; peptide; 7 AA.  
 AC AAY62764;  
 XX 02-MAR-2000 (first entry)  
 DT PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:4047.  
 DE Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT WO9957149-A2.  
 XX

XX 11-NOV-1999.  
 XX 05-MAY-1999; 99WO-CA000363.  
 XX 05-MAY-1999; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX Claim 72; Page 193; 252pp; English.  
 XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug to a tumour through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in  
 CC the exemplification of the present invention  
 XX Sequence 7 AA;  
 SQ  
 Query Match 59.5%; Score 25; DB 3; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 Db 1 CDPKTGC 7  
 RESULT 14  
 AAY62007  
 ID AAY62007 standard; peptide; 7 AA.  
 XX AAY62007;  
 AC AAY62007;  
 XX 02-MAR-2000 (first entry)  
 DT Cadherin-12 cell adhesion recognition cyclic peptide SEQ ID NO:1799.  
 DE Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW neurological disease; cyclic.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT WO9957149-A2.  
 XX

KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT W09957149-A2.  
 XX W09957149-A2.  
 XX 11-NOV-1999.  
 XX 05-MAY-1999; 99WO-CA000363.  
 XX 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX Claim 48; Page 180; 252pp; English.  
 PS The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in  
 CC the exemplification of the present invention  
 XX Sequence 7 AA;  
 SQ Query Match 59.5%; Score 25; DB 3; Length 7;  
 Best Local Similarity 57.1%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 DB 1 CDPKTGC 7  
 RESULT 15  
 AAY62224  
 ID AAY62224 standard; peptide; 7 AA.  
 XX  
 AC AAY62224;  
 XX

DT 02-MAR-2000 (first entry)  
 XX Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:3971.  
 DE Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 XX inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumor; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT W09957149-A2.  
 XX W09957149-A2.  
 XX 11-NOV-1999.  
 XX 05-MAY-1999; 99WO-CA000363.  
 XX 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX Claim 54; Page 184; 252pp; English.  
 PS The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in  
 CC the exemplification of the present invention  
 XX Sequence 7 AA;  
 SQ Query Match 59.5%; Score 25; DB 3; Length 7;  
 Best Local Similarity 57.1%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CVPLTSC 7  
 DB 1 CDPKTGC 7

Db | | | | |  
1 CDPKTC 7

Search completed: September 5, 2004, 11:17:09  
Job time : 64 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 5, 2004, 11:18:40 ; Search time 66 Seconds  
(without alignments)  
33.416 Million cell updates/sec

Title: US-09-761-636A-11

Perfect score: 42

Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 51522

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/FCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	7	9	US-09-761-636A-11
2	31	73.8	7	12	US-09-912-609-7
3	31	73.8	7	12	US-09-922-227-21
4	29	69.0	6	9	US-09-761-636A-12
5	28	66.7	7	10	US-09-932-613-8
6	28	66.7	7	10	US-09-932-322-8
7	26	61.9	7	14	US-10-006-869-1375
8	26	61.9	7	15	US-10-395-032-1375
9	25	59.5	7	14	US-10-006-869-1799
10	25	59.5	7	14	US-10-006-869-3971
11	25	59.5	7	14	US-10-006-869-4047
12	25	59.5	7	15	US-10-395-032-1799
13	25	59.5	7	15	US-10-395-032-3971
14	25	59.5	7	15	US-10-395-032-4047
15	24	57.1	7	14	US-10-006-869-3574

16	24	57.1	7	15	US-10-395-032-3574	Sequence 3574, Ap
17	23	54.8	7	14	US-10-006-869-1392	Sequence 1392, Ap
18	23	54.8	7	14	US-10-006-869-1740	Sequence 1740, Ap
19	23	54.8	7	14	US-10-006-869-2155	Sequence 2155, Ap
20	23	54.8	7	14	US-10-006-869-2720	Sequence 2720, Ap
21	23	54.8	7	14	US-10-006-869-3637	Sequence 3637, Ap
22	23	54.8	7	14	US-10-006-869-4041	Sequence 4041, Ap
23	23	54.8	7	15	US-10-395-032-1392	Sequence 1392, Ap
24	23	54.8	7	15	US-10-395-032-1740	Sequence 1740, Ap
25	23	54.8	7	15	US-10-395-032-2155	Sequence 2155, Ap
26	23	54.8	7	15	US-10-395-032-2720	Sequence 2720, Ap
27	23	54.8	7	15	US-10-395-032-3637	Sequence 3637, Ap
28	23	54.8	7	15	US-10-395-032-4041	Sequence 4041, Ap
29	22	52.4	4	14	US-10-083-894-35	Sequence 221, Appl
30	22	52.4	7	10	US-09-792-286-221	Sequence 225, Appl
31	22	52.4	7	10	US-09-792-286-225	Sequence 1888, Ap
32	22	52.4	7	14	US-10-006-869-1888	Sequence 2750, Ap
33	22	52.4	7	14	US-10-006-869-2750	Sequence 3606, Ap
34	22	52.4	7	14	US-10-006-869-3606	Sequence 3621, Ap
35	22	52.4	7	15	US-10-395-032-1888	Sequence 1888, Ap
36	22	52.4	7	15	US-10-395-032-2750	Sequence 2750, Ap
37	22	52.4	7	15	US-10-395-032-3606	Sequence 3621, Ap
38	22	52.4	7	15	US-10-395-032-3606	Sequence 3606, Ap
39	22	52.4	7	15	US-10-395-032-3621	Sequence 184, Appl
40	21	50.0	6	9	US-09-911-838-184	Sequence 186, Appl
41	21	50.0	6	9	US-09-911-838-186	Sequence 103, Appl
42	21	50.0	7	9	US-09-765-086-103	Sequence 183, Appl
43	21	50.0	7	9	US-09-911-838-183	Sequence 185, Appl
44	21	50.0	7	9	US-09-911-838-185	Sequence 187, Appl
45	21	50.0	7	9	US-09-911-838-187	

## ALIGNMENTS

RESULT 1  
US-09-761-636A-11  
; Sequence 11, Application US/09761636A  
; Patent No. US20020065218A1  
; GENERAL INFORMATION:  
; APPLICANT: ACHEN, Marc  
; APPLICANT: STACKER, Steven  
; APPLICANT: HUGHES, Richard  
; APPLICANT: CENDRON, Angela  
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR  
; FILE REFERENCE: 1064/48505 Achen et al  
; CURRENT APPLICATION NUMBER: US/09/761.636A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/176,293  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/204,590  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-636A-11

Query Match 100.0%; Score 42; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTSC 7  
Db 1 CVPLTSC 7

## RESULT 2

US-09-912-609-7  
; Sequence 7, Application US/09912609  
; Publication No. US20020041898A1

us-09-761-636a-11.closed.rapb

Sun Sep 5 11:41:01 2004

```

; GENERAL INFORMATION:
; APPLICANT: UNGER, EVAN C.
; APPLICANT: MATSUNAGA, TERRY ONICHI
; APPLICANT: RAMANOWSKI, MAREK J.
; APPLICANT: ROMANOWSKI, MAREK J.
; TITLE OF INVENTION: NOVEL TARGETED DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 5030-0001.24
; CURRENT APPLICATION NUMBER: US/09/912,609
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 09/703,474
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/478,124
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-912-609-7

Query Match          73.8%; Score 31; DB 12; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7
Db 1 CLPVASC 7

; RESULT 3
US-09-922-227-21
; Sequence 21, Application US/09922227
; Publication No. US20040071689A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/922,227
; FILING DATE: 02-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; APPLICATION NUMBER: US 09/227,906
; FILING DATE: 08-JAN-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 4859
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-922-227-21

Query Match          73.8%; Score 31; DB 12; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7
Db 1 CLPVASC 7

; RESULT 4
US-09-761-636A-12
; Sequence 12, Application US/09761636A
; Patent No. US20020065218A1
; GENERAL INFORMATION:
; APPLICANT: ACHEN, Marc
; APPLICANT: STACKER, Steven
; APPLICANT: HUGHES, Richard
; APPLICANT: CENDRON, Angela
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR
; FILE REFERENCE: 1064/48505 Achen et al
; CURRENT APPLICATION NUMBER: US/09/761,636A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/176,293
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: US 60/204,590
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-761-636A-12

Query Match          69.0%; Score 29; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPLT 5
Db 1 CVPLT 5

; RESULT 5
US-09-932-613-8
; Sequence 8, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT: DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT

```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Blys binding polypeptide
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X2 is Phe, Trp, or Tyr (preferably Tyr);
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: X4 is Pro or Tyr (preferably Pro);
US-09-932-613-8

Query Match          66.7%; Score 28; DB 10; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CXPXTGC 7

RESULT 6
US-09-932-322-8
; Sequence 8, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Blys binding polypeptide
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X2 is Phe, Trp, or Tyr (preferably Tyr);
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: X4 is Pro or Tyr (preferably Pro);
US-09-932-322-8

Query Match          66.7%; Score 28; DB 10; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CXPXTGC 7

RESULT 7
US-10-006-869-1375
; Sequence 1375, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1375
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1375
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-10-006-869-1375
```

```
Query Match          61.9%; Score 26; DB 14; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CVPLTSC 7
Db 1 CEPKTGC 7
```

```
RESULT 8
US-10-395-032-1375
; Sequence 1375, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1375
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-10-395-032-1375
```

```
Query Match          61.9%; Score 26; DB 15; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CVPLTSC 7
Db 1 CEPKTGC 7
```

```
RESULT 9
US-10-006-869-1799
; Sequence 1799, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1799
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
```

; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence  
US-10-006-869-1799

Query Match 59.5%; Score 25; DB 14; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 10  
US-10-006-869-3971  
; Sequence 3971, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3971  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence  
US-10-006-869-3971

Query Match 59.5%; Score 25; DB 14; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 11  
US-10-006-869-4047  
; Sequence 4047, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4047  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-10-006-869-4047

Query Match 59.5%; Score 25; DB 14; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 12  
US-10-395-032-1799  
; Sequence 1799, Application US/10395032  
; Publication No. US20030229199A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C9  
; CURRENT APPLICATION NUMBER: US/10/395,032  
; CURRENT FILING DATE: 2003-03-21  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1799  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence  
US-10-395-032-1799

Query Match 59.5%; Score 25; DB 15; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 13  
US-10-395-032-3971  
; Sequence 3971, Application US/10395032  
; Publication No. US20030229199A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C9  
; CURRENT APPLICATION NUMBER: US/10/395,032  
; CURRENT FILING DATE: 2003-03-21  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3971  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence  
US-10-395-032-3971

Query Match 59.5%; Score 25; DB 15; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 14  
US-10-395-032-4047

; Sequence 4047, Application US/10395032  
; Publication No. US20030229199A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C9  
; CURRENT APPLICATION NUMBER: US/10/395,032  
; CURRENT FILING DATE: 2003-03-21  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4047  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-10-395-032-4047

Query Match 59.5%; Score 25; DB 15; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
DB 1 CDPKTGC 7

## RESULT 15

US-10-006-869-3574  
; Sequence 3574, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3574  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
; OTHER INFORMATION: recognition sequence  
US-10-006-869-3574

Query Match 57.1%; Score 24; DB 14; Length 7;  
Best Local Similarity 42.9%; Pred. No. 1.2e+06;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
DB 1 CDPVSGC 7

Search completed: September 5, 2004, 11:24:05  
Job time : 67 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:16:10 ; Search time 21 Seconds  
(without alignments)  
17.209 Million cell updates/sec

Title: US-09-761-636A-11  
Perfect score: 42  
Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 57228

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep: \*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep: \*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep: \*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep: \*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep: \*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	73.8	7	1	US-08-526-710-21
2	31	73.8	7	3	US-08-862-855-21
3	31	73.8	7	3	US-09-228-985-21
4	31	73.8	7	4	US-09-227-906-21
5	26	61.9	7	4	US-09-187-859-1375
6	26	61.9	7	4	US-09-839-542B-1375
7	25	59.5	7	4	US-09-187-859-1799
8	25	59.5	7	4	US-09-187-859-3971
9	25	59.5	7	4	US-09-187-859-4047
10	25	59.5	7	4	US-09-839-542B-1799
11	25	59.5	7	4	US-09-839-542B-3971
12	25	59.5	7	4	US-09-839-542B-4047
13	24	57.1	7	4	US-09-187-859-3574
14	24	57.1	7	4	US-09-839-542B-3574
15	23	54.8	6	1	US-07-947-035-13
16	23	54.8	6	1	US-08-321-585A-11
17	23	54.8	7	4	US-09-187-859-1392
18	23	54.8	7	4	US-09-187-859-1740
19	23	54.8	7	4	US-09-187-859-2155
20	23	54.8	7	4	US-09-187-859-2720
21	23	54.8	7	4	US-09-187-859-3637
22	23	54.8	7	4	US-09-187-859-4041
23	23	54.8	7	4	US-09-839-542B-1392
24	23	54.8	7	4	US-09-839-542B-1740
25	23	54.8	7	4	US-09-839-542B-2155
26	23	54.8	7	4	US-09-839-542B-2720
27	23	54.8	7	4	US-09-839-542B-3637

28	23	54.8	7	4	US-09-839-542B-4041	Sequence 4041, Ap
29	23	54.8	7	4	US-09-535-852-1826	Sequence 1826, Ap
30	22	52.4	4	2	US-08-429-964-41	Sequence 41, Appl
31	22	52.4	4	2	US-09-665-362A-35	Sequence 35, Appl
32	22	52.4	4	5	PCT-US93-08062-41	Sequence 41, Appl
33	22	52.4	7	1	US-07-958-903A-26	Sequence 26, Appl
34	22	52.4	7	1	US-08-462-018-26	Sequence 26, Appl
35	22	52.4	7	1	US-08-823-245-26	Sequence 26, Appl
36	22	52.4	7	4	US-07-963-329A-26	Sequence 26, Appl
37	22	52.4	7	4	US-09-187-859-1888	Sequence 1888, Ap
38	22	52.4	7	4	US-09-187-859-2750	Sequence 2750, Ap
39	22	52.4	7	4	US-09-187-859-3606	Sequence 3606, Ap
40	22	52.4	7	4	US-09-187-859-3621	Sequence 3621, Ap
41	22	52.4	7	4	US-09-839-542B-1888	Sequence 1888, Ap
42	22	52.4	7	4	US-09-839-542B-2750	Sequence 2750, Ap
43	22	52.4	7	4	US-09-839-542B-3606	Sequence 3606, Ap
44	22	52.4	7	4	US-09-839-542B-3621	Sequence 3621, Ap
45	22	52.4	7	4	US-09-535-852-918	Sequence 918, App

## ALIGNMENTS

RESULT 1  
US-08-526-710-21  
; Sequence 21, Application US/08526710  
; Patent No. 5622699  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; TITLE OF INVENTION: Method of Identifying Molecules That  
; TITLE OF INVENTION: Home to a Selected Organ In Vivo  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/526,710  
; FILING DATE: 11-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-526-710-21

Query Match 73.8%; Score 31; DB 1; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CVPLTSC 7  
|:|:|  
Db 1 CLPVASC 7

us-09-761-636a-11.closed.ra1

Sun Sep 5 11:41:01 2004

```
RESULT 2
US-08-862-855-21
; Sequence 21, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-862-855-21

Query Match 73.8%; Score 31; DB 3; Length 7;
Best Local Similarity 57.1%; Pred. NO. 3e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CLPVASC 7

RESULT 3
US-09-226-985-21
; Sequence 21, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-862-855-21

Query Match 73.8%; Score 31; DB 3; Length 7;
Best Local Similarity 57.1%; Pred. NO. 3e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CLPVASC 7

RESULT 4
US-09-227-906-21
; Sequence 21, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-226-985-21

Query Match 73.8%; Score 31; DB 3; Length 7;
Best Local Similarity 57.1%; Pred. NO. 3e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CLPVASC 7
```





Query Match 59.5%; Score 25; DB 4; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 9  
US-09-187-859-4047  
; Sequence 4047, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 4047  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-4047

Query Match 59.5%; Score 25; DB 4; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 10  
US-09-839-542B-1799  
; Sequence 1799, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1799  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence  
US-09-839-542B-1799

Query Match 59.5%; Score 25; DB 4; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 11  
US-09-839-542B-3971  
; Sequence 3971, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3971  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence  
US-09-839-542B-3971

Query Match 59.5%; Score 25; DB 4; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 12  
US-09-839-542B-4047  
; Sequence 4047, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 4047  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542B-4047

Query Match 59.5%; Score 25; DB 4; Length 7;  
Best Local Similarity 57.1%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPKTGC 7

RESULT 13  
US-09-187-859-3574  
; Sequence 3574, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.  
APPLICANT: Gour, Barbara J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
FILE REFERENCE: 100086.407C1  
CURRENT APPLICATION NUMBER: US/09/187,859A  
CURRENT FILING DATE: 1998-11-06  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3574  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
OTHER INFORMATION: recognition sequence

US-09-187-859-3574

Query Match 57.1%; Score 24; DB 4; Length 7;  
Best Local Similarity 42.9%; Pred. No. 3e+05; 2; Mismatches 2; Indels 0; Gaps 0;  
Matches 3; Conservative 2;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPVSGC 7

RESULT 14  
US-09-839-542B-3574  
Sequence 3574, Application US/09839542B  
Patent No. 6569996  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.  
APPLICANT: Symonds, James Matthew  
APPLICANT: Gour, Barbara J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
FILE REFERENCE: 100086.407D1  
CURRENT APPLICATION NUMBER: US/09/839,542B  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3574  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
OTHER INFORMATION: recognition sequence

US-09-839-542B-3574

Query Match 57.1%; Score 24; DB 4; Length 7;  
Best Local Similarity 42.9%; Pred. No. 3e+05; 2; Mismatches 2; Indels 0; Gaps 0;  
Matches 3; Conservative 2;

QY 1 CVPLTSC 7  
| | | | |  
Db 1 CDPVSGC 7

RESULT 15  
US-07-947-035-13  
Sequence 13, Application US/07947035  
Patent No. 544045  
GENERAL INFORMATION:  
APPLICANT: Francis, Geoffrey L.  
APPLICANT: Walton, Paul E.  
APPLICANT: Ballard, Francis J.  
APPLICANT: McMurty, John P.  
APPLICANT: Phelps, Patricia V.  
TITLE OF INVENTION: Method of Administering IGF-1, IGF-2,

TITLE OF INVENTION: and Analogs Thereof to Birds  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenneth D. Sibley  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 5444045th Carolina  
COUNTRY: US  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/947,035  
FILING DATE: 17-SEP-1992  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sibley, Kenneth D.  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5175-59  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 881-3140  
TELEFAX: (919) 881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-07-947-035-13

Query Match 54.8%; Score 23; DB 1; Length 6;  
Best Local Similarity 80.0%; Pred. No. 3e+05; 0; Mismatches 1; Indels 0; Gaps 0;  
Matches 4; Conservative 0;

QY 3 PLTSC 7  
| | | | |  
Db 2 PLTLC 6

Search completed: September 5, 2004, 11:19:28  
Job time : 22 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:22:56 ; Search time 21 seconds  
(without alignments)  
27.483 Million cell updates/sec

Title: US-09-761-636A-12

Perfect score: 38

Sequence: 1 CVPLTC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 316

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	52.6	5	2 B22565	R-phycoerythrin al
2	15	39.5	6	2 I65546	MHC H2-L antigen
3	14	36.8	5	2 E60274	major protein anti
4	14	36.8	5	2 F22565	R-phycoerythrin ga
5	14	36.8	6	2 I49421	laminin B1 - waste
6	11	28.9	4	2 A32039	tyrosine-melanocyt
7	11	28.9	4	2 I54357	schwannomin - mous
8	11	28.9	5	2 A60521	glycogen phosphory
9	11	28.9	6	2 I37263	Y protein - human
10	11	28.9	6	2 H48394	glycoprotein compo
11	10	26.3	3	2 C22565	R-phycoerythrin be
12	9	23.7	3	3 A22565	R-phycoerythrin al
13	9	23.7	4	2 I51049	metallothionein-A
14	9	23.7	4	2 S43959	Ig mu chain V regi
15	9	23.7	4	2 S55238	pallidipin - assas
16	9	23.7	5	2 A33882	cadmium-binding pe
17	9	23.7	5	2 B37988	acid proteinase li
18	9	23.7	5	2 B45525	actin I - malaria
19	9	23.7	5	2 S65726	hemoglobin, extrac
20	9	23.7	5	2 S11127	phosphoprotein, bo
21	9	23.7	6	2 A61049	halo-toxin - Pseud
22	9	23.7	6	2 J00355	lipopeptide Ws1279
23	9	23.7	6	2 P00008	angiotensin-conver
24	9	23.7	6	2 A27696	contraction-inhibi
25	9	23.7	6	2 B27696	contraction-inhibi
26	9	23.7	6	2 I37027	protamine P1 - gor
27	9	23.7	6	2 I67345	MHC H2-K-k cell su
28	9	23.7	6	2 P07052	T-cell receptor be
29	9	23.7	6	2 F41946	T-cell receptor ga

30 9 23.7 6 2 S29881 Na+/K+-exchanging  
31 9 23.7 6 4 I79564 hypothetical TCL3  
32 8 21.1 3 1 I78890 tyrosine protein k  
33 8 21.1 5 1 HOROHA proctolin - Americ  
34 8 21.1 5 2 JN0862 peptidyl-dipeptida  
35 8 21.1 5 2 JN0860 peptidyl-dipeptida  
36 8 21.1 5 2 E42364 flagellar protein  
37 8 21.1 5 2 P00009 angiotensin-conver  
38 8 21.1 5 2 A60411 proctolin - Atlant  
39 8 21.1 5 2 S53595 hypothetical prote  
40 8 21.1 6 4 A35039 hypothetical colla  
41 7 18.4 3 3 RHTDTC thyrroliberin - Bom  
42 7 18.4 3 3 RHPGT thyrroliberin - pig  
43 7 18.4 3 3 RSHST thyrroliberin - she  
44 7 18.4 3 3 A92971 thyrroliberin - eas  
45 7 18.4 3 3 A33802 thyrotropin-releas

## ALIGNMENTS

### RESULT 1

B22565

R-phycoerythrin alpha-2 chain - red alga (Gastrocloonium coulteri) (fragment)

C:Species: Gastrocloonium coulteri

C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993

C:Accession: B22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A:Reference number: A22565; MUID:85182601; PMID:3886644

A:Accession: B22565

A:Molecule type: protein

A:Residues: 1-5 <KLO>

Query Match 52.6%; Score 20; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVP 3

Db 2 CVP 4

### RESULT 2

I65546

MHC H2-L antigen - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999

C:Accession: I65546

R:Kimura, A.; Israel, A.; Le Bail, O.; Kourilsky, P.

Cell 44, 261-272, 1986

A:Title: Detailed analysis of the mouse H-2Kb promoter: Enhancer-like sequences and their

A:Reference number: I52778; MUID:86106202; PMID:3510743

A:Accession: I65546

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-6 <RES>

A:Cross-references: GB:M12483; NID:gi99565; PIDN:AAA39663.1; PID:gs54234

Query Match

Best Local Similarity 39.5%; Score 15; DB 2; Length 6;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VPLT 5

Db 2 VPCT 5

### RESULT 3

E60274

major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)

C:Species: Mycobacterium tuberculosis

C>Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 30-Sep-1993  
 C:Accession: E60274  
 R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.  
 Infect. Immun. 59, 372-382, 1991  
 A>Title: Isolation and partial characterization of major protein antigens in the culture  
 A:Reference number: A60274; MUID:9109989; PMID:1898899  
 A:Accession: E60274  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-5 <NAG>

Query Match 36.8%; Score 14; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLT 5  
 ||  
 Db 3 PIT 5

## RESULT 4

F22565  
 R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)  
 C:Species: Gastroclonium coulteri  
 C>Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993  
 C:Accession: F22565  
 R:Klotz, A.V.; Glazer, A.N.  
 J. Biol. Chem. 260, 4856-4863, 1985  
 A>Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
 A:Reference number: A22565; MUID:85182601; PMID:3886644  
 A:Accession: F22565  
 A:Molecule type: protein  
 A:Residues: 1-5 <KLO>

Query Match 36.8%; Score 14; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TC 6  
 ||  
 Db 2 TC 3

## RESULT 5

I49421  
 laminin B1 - western wild mouse (fragment)  
 C:Species: Mus spretus (western wild mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C:Accession: I49421  
 R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.  
 Mamm. Genome 5, 349-355, 1994  
 A>Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
 A:Reference number: I48934; MUID:94319082; PMID:8043949  
 A:Accession: I49421  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-6 <RES>  
 A:Cross-references: EMBL:U05736; NID:G497073; PIDN:AB60477.1; PID:G642829

Query Match 36.8%; Score 14; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TC 6  
 ||  
 Db 4 TC 5

## RESULT 6

A32039  
 tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 31-Jul-1989 #sequence\_revision 31-Jul-1989 #text\_change 18-Aug-2000

C:Accession: A32039  
 R:Horvath, A.; Kastin, A.J.  
 J. Biol. Chem. 264, 2175-2179, 1989  
 A>Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor  
 A:Reference number: A32039; MUID:89123285; PMID:2563371  
 A:Accession: A32039  
 A:Molecule type: protein  
 A:Residues: 1-4 <HOR>

A:Experimental source: brain  
 A:Superfamily: unassigned animal peptides  
 C:Keywords: amidated carboxyl end  
 F:4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.9%; Score 11; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4  
 ||  
 Db 2 PL 3

## RESULT 7

I54357  
 schwannomin - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C:Accession: I54357  
 R:Huyhn, D.P.; Nechiporuk, T.; Pulst, S.  
 Hum. Mol. Genet. 3, 1075-1079, 1994  
 A>Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are co  
 A:Reference number: I54357; MUID:95072570; PMID:7981675  
 A:Accession: I54357  
 A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA  
 A:Residues: 1-4 <RES>  
 A:Cross-references: GB:I28838; NID:G454836; PIDN:AAA57150.1; PID:G601923  
 C:Genetics:  
 A:Gene: NF2

Query Match 28.9%; Score 11; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3  
 ||  
 Db 1 VP 2

## RESULT 8

A60521  
 glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)  
 N:Alternate names: glycogen phosphorylase b  
 C:Species: Liza ramada  
 C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2003  
 C:Accession: A60521  
 R:Bonamusa, L.; Baanante, I.V.  
 Comp. Biochem. Physiol. B 95, 295-301, 1990  
 A>Title: Purification and characterization of glycogen phosphorylase B from skeletal mus  
 A:Reference number: A60521; MUID:90227907; PMID:2109669  
 A:Accession: A60521  
 A:Molecule type: protein  
 A:Residues: 1-5 <BON>

C:Superfamily: glucan phosphorylase  
 C:Keywords: glycocyltransferase; hexosyltransferase; phosphoprotein  
 F:3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experim

Query Match 28.9%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3  
 ||

Db 4 VP 5

RESULT 9  
I37263  
Y protein - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I37263  
R:Waerber, G.; Habener, J.F.  
Endocrinology 131, 2010-2015, 1992  
A>Title: Novel testis germ cell-specific transcript of the CREB gene contains an alternative exon  
A:Reference number: I37263; MUID:93010691; PMID:11396344  
A:Accession: I37263  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: EMBL:X68994; NID:9396171; PIDN:CAA48780.1; PID:G579816  
A:Gene: CREB

Query Match  
Best Local Similarity 28.9%; Score 11; DB 2; Length 6;  
Best Local Similarity 56.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTC 6  
|  
|  
|  
Db 2 LTC 4

RESULT 10  
H48394  
Glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C:Accession: H48394  
R:Maether, I.H.; Banghart, L.R.; Lane, W.S.  
Biochem. Mol. Biol. Int. 29, 545-554, 1993  
A>Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.  
A:Reference number: A48394; MUID:93250576; PMID:8485470  
A:Accession: H48394  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <MAT>  
A:Experimental source: milk  
A>Note: sequence extracted from NCBI backbone (NCBIP:131518)  
C:Keywords: glycoprotein

Query Match  
Best Local Similarity 28.9%; Score 11; DB 2; Length 6;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTC 6  
|  
|  
|  
Db 4 LTC 6

RESULT 11  
C22565  
R-phycoerythrin beta-1 chain - red alga (Gastroclonium coulteri) (fragment)  
C:Species: Gastroclonium coulteri  
C>Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993  
C:Accession: C22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A>Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: C22565  
A:Molecule type: protein  
A:Residues: 1-6 <KLO>

Query Match  
Best Local Similarity 26.3%; Score 10; DB 2; Length 6;

Db 4 VP 5

RESULT 10  
I37263  
Y protein - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I37263  
R:Waerber, G.; Habener, J.F.  
Endocrinology 131, 2010-2015, 1992  
A>Title: Novel testis germ cell-specific transcript of the CREB gene contains an alternative  
A:Reference number: I37263; MUID:93010691; PMID:11396344  
A:Accession: I37263  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: EMBL:X68994; NID:9396171; PIDN:CAA48780.1; PID:G579816  
A:Gene: CREB

Query Match  
Best Local Similarity 28.9%; Score 11; DB 2; Length 6;  
Best Local Similarity 56.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTC 6  
|  
|  
|  
Db 2 LTC 4

RESULT 10  
H48394  
Glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C:Accession: H48394  
R:Maether, I.H.; Banghart, L.R.; Lane, W.S.  
Biochem. Mol. Biol. Int. 29, 545-554, 1993  
A>Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.  
A:Reference number: A48394; MUID:93250576; PMID:8485470  
A:Accession: H48394  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <MAT>  
A:Experimental source: milk  
A>Note: sequence extracted from NCBI backbone (NCBIP:131518)  
C:Keywords: glycoprotein

Query Match  
Best Local Similarity 28.9%; Score 11; DB 2; Length 6;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LTC 6  
|  
|  
|  
Db 4 LTC 6

RESULT 11  
C22565  
R-phycoerythrin beta-1 chain - red alga (Gastroclonium coulteri) (fragment)  
C:Species: Gastroclonium coulteri  
C>Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993  
C:Accession: C22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A>Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: C22565  
A:Molecule type: protein  
A:Residues: 1-6 <KLO>

Query Match  
Best Local Similarity 26.3%; Score 10; DB 2; Length 6;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CV 2  
|  
|  
|  
Db 4 CL 5

## RESULT 12

A22565  
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)  
C:Species: Gastroclonium coulteri  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A>Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: A22565  
A:Molecule type: protein  
A:Residues: 1-3 <KLO>

Query Match  
Best Local Similarity 23.7%; Score 9; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 C 1  
|  
|  
|  
Db 1 C 1

## RESULT 13

I51049  
Metallothionein-A - rainbow trout (fragment)  
C:Species: Oncorhynchus mykiss (rainbow trout)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I51049  
R:Olsson, P.E.; Kling, P.; Erkel, L.J.; Kille, P.  
Eur. J. Biochem. 230, 344-349, 1995  
A>Title: Structural and functional analysis of the rainbow trout (Oncorhynchus mykiss) me  
A:Reference number: I51049; MUID:95324545; PMID:7601121  
A:Accession: I51049  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <OLS>  
A:Cross-references: EMBL:X80181; NID:G1019799; PIDN:CAA56466.1; PID:G4379328

Query Match  
Best Local Similarity 23.7%; Score 9; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 C 1  
|  
|  
|  
Db 4 C 4

## RESULT 14

S43959  
Ig mu chain V region (clone 13) - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 20-Oct-1994 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999  
C:Accession: S43959  
R:Wagner, S.D.; Williams, G.T.; Larson, T.; Neuberger, M.S.; Kitamura, D.; Rajewsky, K.;  
Nucleic Acids Res. 22, 1389-1393, 1994  
A>Title: Antibodies generated from human immunoglobulin miniloci in transgenic mice.  
A:Reference number: S43956; MUID:94248036; PMID:8190629  
A:Accession: S43959  
A:Molecule type: DNA  
A:Residues: 1-4 <WAG>  
C:Keywords: immunoglobulin

Query Match  
Best Local Similarity 23.7%; Score 9; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

```
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 C 1
      |
Db      2 C 2

RESULT 15
S55238
pallidipin - assassin bug (fragment)
C:Species: Triatoma pallidipennis (assassin bug)
C>Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 19-May-2000
C:Accession: S55238
R:Haendler, B.; Becker, A.; Noeske-Jungblut, C.; Kraetzschmar, J.; Donner, P.; Schleunin
Biochem. J. 307, 465-470, 1995
A:Title: Expression of active recombinant pallidipin, a novel platelet aggregation inhib
A:Reference number: S55238; MUID:95251610; PMID:7733884
A:Accession: S55238
A:Molecule type: protein
A:Residues: 1-4 <HAE>

Query Match      23.7%; Score 9; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 C 1
      |
Db      3 C 3

Search completed: September 5, 2004, 11:27:07
Job time : 21 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:19:31; Search time 15 Seconds  
(without alignments)  
20.828 Million cell updates/sec

Title: US-09-761-636A-12

Perfect score: 28

Sequence: 1 CVPLTC 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 50

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	13	34.2	6	EI01_LITRU	P82096 litoria rub
2	11	28.9	6	VP19_HSV1X	P23210 herpes slimp
3	9	23.7	5	EI03_LITRU	P82099 litoria rub
4	9	23.7	6	CIP1_MYTED	P13736 mytilus edu
5	9	23.7	6	CIP2_MYTED	P13737 mytilus edu
6	8	21.1	5	PRCT_PERAM	P01373 periplaneta
7	8	21.1	6	TRP1_PSEPU	P36414 pseudomonas
8	7	18.4	3	THYL_PIG	P01151 sus scrofa
9	7	18.4	4	DCML_PSECH	P19916 pseudomonas
10	7	18.4	4	RM01_YEAST	P36515 saccharomyc
11	7	18.4	4	TUFT_HUMAN	P01858 homo sapien
12	7	18.4	5	BIOA_CITFR	P13071 citrobacter
13	7	18.4	5	BPF7_BOTIN	P30425 bothrops in
14	7	18.4	5	EI04_LITRU	P82100 litoria rub
15	7	18.4	5	PAF2_PARMA	P81864 pardachirus
16	7	18.4	5	SUGA_ACHDO	P19991 acheta dome
17	7	18.4	6	OMV_LRPDE	P42985 leptonotars
18	7	18.4	6	TWOF_SARBU	P41495 sarcophaga
19	5	13.2	5	PSK_DAUCA	P58261 daucus caro
20	5	13.2	5	R331_LITRU	P82072 litoria rub
21	5	13.2	5	R332_LITRU	P82073 litoria rub
22	5	13.2	6	UN06_CLOPA	P81351 clostridium
23	4	10.5	4	EOSI_HUMAN	P02731 homo sapien
24	4	10.5	4	FAR3_HIRME	P42562 hirudo medi
25	4	10.5	4	ELRF_HIRME	P42561 hirudo medi
26	4	10.5	4	FLRN_ATEL	P58707 anthopleura
27	4	10.5	5	RE11_LITRU	P81817 carcinus ma
28	4	10.5	5	AL14_CARNA	P82070 litoria rub
29	4	10.5	5	TPIS_CANFA	P54714 canis fami
30	4	10.5	5	UC22_MAIZE	P80628 zea mays (m
31	4	10.5	6	ACPH_RABIT	P25154 oryctolagus
32	3	7.9	3	LUXE_VIBFI	P24272 vibrio fisc
33	3	7.9	4	FYRI_ATEL	P58706 anthopleura

34 3 7.9 5 1 FARP\_ARTTR P41853 artiopesthi  
35 3 7.9 5 1 RE21\_LITRU P82071 litoria rub  
36 3 7.9 5 1 UF01\_MOUSE P38639 mus musculus  
37 2 5.3 4 1 DCMS\_PSECH P19918 pseudomonas  
38 2 5.3 4 1 FAR4\_HIRME P42563 hirudo medi  
39 2 5.3 4 1 FMRF\_MAGNI P01162 macrocallis  
40 2 5.3 5 1 BIOB\_CITFR P12997 citrobacter  
41 1 2.6 4 1 OCP3\_OCTMI P58649 octopus min  
42 1 2.6 5 1 UXA4\_CHLTR P38005 chlamydia t  
43 1 2.6 6 1 LOK1\_LOCFI P41491 locusta mig  
44 0 0.0 3 1 GRWM\_HUMAN P01157 homo sapien  
45 0 0.0 4 1 ACHJ\_ACHFU P35904 achatina fu

## ALIGNMENTS

RESULT 1  
EI01\_LITRU STANDARD; PRT; 6 AA.  
AC P82096;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Electrin 1.  
OS Litoria rubella (Desert tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Hylidae;  
OC Pelodyadinae; Litoria.  
OX NCBI\_TaxID=104895;  
RN [1] SEQUENCE.  
RP TISSUE=Skin secretion;  
RC Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
RT "Peptides from the skin glands of the Australian buzzing tree frog  
RT Litoria electrica. Comparison with the skin peptides from Litoria  
RT rubella.";  
RL Aust. J. Chem. 52:639-645(1999).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Skin.  
KW Amphibian defense peptide; Amidation.  
FT MOD RES 6  
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 34.2%; Score 13; DB 1; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.4e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPL 4  
DB 2 VPI 4

RESULT 2  
VP19\_HSV1X STANDARD; PRT; 6 AA.  
ID VP19\_HSV1X  
AC P23210;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Capsid assembly and DNA maturation protein (Virion protein UL38)  
DE (Capsid protein VP19C) (Fragment).  
GN UL38.  
OS Herpes simplex virus (type 1 / strain KOS).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10306;  
RN [1] SEQUENCE FROM N.A.  
RP MEDLINE=91101287; PubMed=1846198;  
RX Flanagan W.M., Papavassiliou A.G., Rice M., Hecht L.B.,  
RA Silverstein S., Wagner E.K.;  
RT "Analysis of the herpes simplex virus type 1 promoter controlling the

Sun Sep 5 11:41:05 2004

us-09-761-636a-12.closed.rsp

```

RT expression of UL38, a true late gene involved in capsid assembly.";
RL J. Virol. 65:769-786(1991).
CC -!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -!- SIMILARITY: Belongs to the herpesviruses VP19C family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M57646; AAA45830.1; -
CC Capsid assembly; Coat protein; DNA-binding.
CC NON_TER 6
CC SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;
DR EMBL; M57646; AAA45830.1; -
KW Capsid assembly; Coat protein; DNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 28.9%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 5 PL 6

RESULT 3
EIO3_LITRU STANDARD; PRT; 5 AA.
AC P82059;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 3.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
CC Pelodyadinae; Litoria.
CC NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella".
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 5
SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 4 PL 5

RESULT 4
CIP1_MYTED STANDARD; PRT; 6 AA.
AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide I (MIP I).
OS Mytilus edulis (Blue mussel).

expression of UL38, a true late gene involved in capsid assembly.";
J. Virol. 65:769-786(1991).
-!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
EMBEDDED. BINDS DNA.
-!- SIMILARITY: Belongs to the herpesviruses VP19C family.
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL; M57646; AAA45830.1; -
Capsid assembly; Coat protein; DNA-binding.
NON_TER 6
SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;
DR EMBL; M57646; AAA45830.1; -
KW Capsid assembly; Coat protein; DNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 28.9%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 5 PL 6

RESULT 3
EIO3_LITRU STANDARD; PRT; 5 AA.
AC P82059;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 3.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
CC Pelodyadinae; Litoria.
CC NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella".
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 5
SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 4 PL 5

RESULT 4
CIP1_MYTED STANDARD; PRT; 6 AA.
AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide I (MIP I).
OS Mytilus edulis (Blue mussel).

```

```

CC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
CC Mytiloidea; Mytilidae; Mytilus.
CC NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE=Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -!- FUNCTION: Inhibitory action on contractions in several molluscan
CC muscles.
CC -!- SIMILARITY: TO MIP II.
DR PIR; A27696; A27696.
KW Hormone; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 637 MW; 72C9C68775B81000 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 3 PM 4

RESULT 5
CIP2_MYTED STANDARD; PRT; 6 AA.
AC P13737;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide II (MIP II).
OS Mytilus edulis (Blue mussel).
CC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
CC Mytiloidea; Mytilidae; Mytilus.
CC NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE=Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -!- FUNCTION: Inhibitory action on contractions in several molluscan
CC muscles.
CC -!- SIMILARITY: TO MIP I.
DR PIR; B27696; B27696.
KW Hormone; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4
DB 3 PM 4

RESULT 6
PRCT_PERAM STANDARD; PRT; 5 AA.
AC P01373;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Proctolin.
OS Periplaneta americana (American cockroach),
OS Limulus polyphemus (Atlantic horseshoe crab), and

```

OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;  
 OC Blattidae; Periplaneta.  
 OX NCBI\_TaxID=6978, 6850, 6759;  
 RN [1]  
 RP SEQUENCE  
 RC SPECIES=P.americana;  
 RX MEDLINE=76074708; PubMed=576;  
 RA Starratt A.N., Brown B.E.;  
 RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter  
 in insects";  
 RL Life Sci. 17:1253-1256(1975).  
 RN [2]  
 RP BIOLOGICAL SOURCE  
 RC SPECIES=P.americana;  
 RX MEDLINE=81225865; PubMed=6113690;  
 RA O'Shea M., Adams M.E.;  
 RT "Pentapeptide (proctolin) associated with an identified neuron";  
 RL Science 213:567-569(1981).  
 RN [3]  
 RP SEQUENCE  
 RC SPECIES=L.polyphemus;  
 RX MEDLINE=90287800; PubMed=2356151;  
 RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,  
 RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,  
 RA Shabanowitz J.;  
 RT "Identification of proctolin in the central nervous system of the  
 horseshoe crab, Limulus polyphemus";  
 RL Peptides 11:205-211(1990).  
 RN [4]  
 RP SEQUENCE  
 RC SPECIES=C.maenas;  
 RX MEDLINE=86232789; PubMed=2872661;  
 RA Stangier J., Dirksen H., Keller R.;  
 RT "Identification and immunocytochemical localization of proctolin in  
 pericardial organs of the shore crab, Carcinus maenas";  
 RL Peptides 7:67-72(1986).  
 CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,  
 modulates visceral and skeletal muscle in many arthropods.  
 CC -!- TISSUE SPECIFICITY: Found in the lateral white neurons and in  
 the crab pericardial organs.  
 CC PIR; A01644; HOROHA.  
 DR PIR; A60411; A60411.  
 KW Neuroptide.  
 SQ SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;  
 Query Match 21.1%; Score 8; DB 1; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VP 3  
 Db :  
 3 LP 4  
 RESULT 7  
 TRPI\_PSEPU  
 ID TRPI\_PSEPU STANDARD; PRT; 6 AA.  
 AC P36414;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE HTH-type transcriptional regulator trpI (TrpBA operon transcriptional  
 activator) (Fragment).  
 GN TRPI.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PPG1 C1S;  
 RX MEDLINE=89335826; PubMed=2503057;  
 RA Eberly L., Crawford I.P.;  
 RT "DNA sequence of the tryptophan synthase genes of Pseudomonas  
 putida";  
 RL Biochimie 71:521-531(1989).  
 CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPBA GENES ENCODING  
 THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE  
 INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE  
 TRPAB OPERON, OVERLAPPING ITS OWN PROMOTER REGION.  
 CC -!- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X13299; CAA31660.1; -  
 DR InterPro; IPR000847; HTH\_LysR.  
 DR PROSITE; PS50931; HTH\_LYSR; PARTIAL.  
 KW Tryptophan biosynthesis; Transcription regulation; Activator;  
 KW DNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 683 MW; 77672AA1EDD6F000 CRC64;  
 Query Match 21.1%; Score 8; DB 1; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VP 3  
 Db :  
 5 LP 6  
 RESULT 8  
 THYL\_PIG  
 ID THYL\_PIG STANDARD; PRT; 3 AA.  
 AC P01151;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).  
 OS Sus scrofa (Pig),  
 OS Sus aries (Sheep),  
 OS Bombina orientalis (Oriental fire-bellied toad), and  
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823, 9940, 8346, 8316;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=Pig; TISSUE=Hypothalamus;  
 RX MEDLINE=70136150; PubMed=4984938;  
 RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
 RT "Structure of porcine thyrotropin releasing hormone";  
 RL Biochemistry 9:1103-1106(1970).  
 RN [2]  
 RP SYNTHESIS.  
 RC SPECIES=Pig;  
 RX MEDLINE=70039904; PubMed=4982117;  
 RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
 RT "The identity of chemical and hormonal properties of the thyrotropin  
 releasing hormone and pyroglutamyl-histidyl-proline amide";  
 RL Biochem. Biophys. Res. Commun. 37:705-710(1969).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=Sheep; TISSUE=Hypothalamus;  
 RA Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,  
 RA Ward D.N.;  
 RT "The elucidation of the primary structure of the hypothalamic thyroid  
 stimulating hormone releasing factor of ovine origin by means of mass

```

RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228 (1971).
RN [4]
RP SYNTHESIS.
RC SPECIES=Sheep; PubMed=4985794;
RX MEDLINE=70163386; PubMed=4985794;
RA Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RT Guillemain R.;
RA "Characterization of ovine hypothalamic hypophysiotropic
RT TSH-releasing factor.";
RL Nature 226:321-325 (1970).
RN [5]
RP SEQUENCE.
RC SPECIES=B. orientalis; TISSUE=Skin;
RX MEDLINE=76138399; PubMed=815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RN Chem. Pharm. Bull. 23:3301-3303 (1975).
RN [6]
RP SEQUENCE.
RC SPECIES=N. viridescens;
RX MEDLINE=75035605; PubMed=4214528;
RA Grimm-Joergensen Y., McKelvy J.P.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478 (1974).
RC J. FUNCTION: TRH functions as a regulator of the biosynthesis of TSH
in the anterior pituitary gland and as a neurotransmitter/
neuromodulator in the central and peripheral nervous systems.
CC PIR; A90919; RHDTIO.
DR PIR; A92971; A92971.
DR PIR; A93750; RSHST.
KW Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 3 3 AMIDATION.
SQ SEQUENCE 3 AA; 380 MW; 776196B000000000 CRC64;

Query Match 18.4%; Score 7; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 P 3
Db 3 P 3

RESULT 9
DCML_PSECH STANDARD; PRT; 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GN CUTL.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydohydrophic bacteria.";
RL Arch. Microbiol. 152:335-341 (1989).
CC Arch. FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -I- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -I- COFACTOR: Molybdenum (molybdopterin).

```

---

```

CC -I- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; PLO140; PLO140.
KW Oxidoreductase; Molybdenum.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 441 MW; 7761E876F000000000 CRC64;

Query Match 18.4%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 P 3
Db 4 P 4

RESULT 10
RM01 YEAST STANDARD; PRT; 4 AA.
AC P36515;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
GN MRPL1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RX MEDLINE=91285106; PubMed=2060626;
RA Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
RT "Extended N-terminal sequencing of proteins of the large ribosomal
RT subunit from yeast mitochondria.";
RL FEBS Lett. 284:51-56 (1991).
DR PIR; S17255; S17255.
DR SGD; L0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D000000000 CRC64;

Query Match 18.4%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 P 3
Db 4 P 4

RESULT 11
TUFT HUMAN STANDARD; PRT; 4 AA.
AC P01858;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phagocytosis-stimulating peptide (Tuftsin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishioka K., Constantopoulos A., Satch P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
RT stimulating peptide tuftsin.";
RL Biochem. Biophys. Res. Commun. 47:172-179 (1972).
RN [2]
RP IMMUNOGLOBULIN CLASS.

```

RX MEDLINE=68091045; PubMed=4169272;  
RA Fidalgo B.V., Najjar V.A.;  
RT "The physiological role of the lymphoid system. VI. The stimulatory  
RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic  
RT activity of human polymorphonuclear leucocyte.";  
RL Biochemistry 6:3386-3392(1967).  
CC -!- MISCELLANEOUS: An IGG (called leucokinin) binds reversibly to the  
CC cell membrane of neutrophils in the blood. Leucokininase on the  
CC membrane releases the active peptide tuftsin from the gamma chain.  
CC Tuftsin is essential for maximum stimulation of the phagocytic  
CC activity of neutrophils.  
DR PIR; A02147; A02147.  
DR MIM; 191150; -.  
DR GO; GO:0003823; P:antigen binding; NAS.  
DR GO; GO:0006909; P:phagocytosis; NAS.  
SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;  
  
Query Match 18.4%; Score 7; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 P 3  
Db 3 P 3  
  
RESULT 12  
BIOA CITR STANDARD; PRT; 5 AA.  
AC P13071;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase  
DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA  
DE aminotransferase) (Fragment).  
GN BIOA.  
OS Citrobacter freundii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Citrobacter.  
OX NCBI\_TaxID=546;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89006280; PubMed=2971595;  
RA Shuan D., Campbell A.;  
RT "Transcriptional regulation and gene arrangement of Escherichia coli,  
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";  
RL Gene 67:203-211(1988).  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-  
CC oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-  
CC diaminononanoate.  
CC -!- COFACTOR: Pyridoxal phosphate.  
CC -!- PATHWAY: Biotin biosynthesis.  
CC -!- SUBUNIT: Homodimer.  
CC -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent  
CC aminotransferases.  
  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
  
EMBL; M21922; -; NOT\_ANNOTATED CDS.  
DR PIR; I40697; I40697.  
DR InterPro; IPR005814; AminoTrans 3.  
DR PROSITE; PS00600; AA\_TRANSFCLASS 3; PARTIAL.  
KW Biotin biosynthesis; Transferase; Aminotransferase;  
KW Pyridoxal phosphate.  
FT NON TER 5  
SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;

Query Match 18.4%; Score 7; DB 1; Length 5;  
Best Local Similarity 50.0%; Pred. No. 1.4e+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 LT 5  
Db 1 MT 2  
  
RESULT 13  
BPP7 BOTIN STANDARD; PRT; 5 AA.  
AC P30425;  
DT 01-FEB-1993 (Rel. 25, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Bradykinin-potentiating peptide S5,2 (5A) (Angiotensin-converting  
DE enzyme inhibitor).  
OS Bothrops insularis (Island jararaca) (Queimada jararaca).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Crotalinae; Bothrops.  
OX NCBI\_TaxID=8723;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=90351557; PubMed=2386615;  
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;  
RT "Primary structure and biological activity of bradykinin potentiating  
RT peptides from Bothrops insularis snake venom.";  
RL J. Protein Chem. 9:221-227(1990).  
CC -!- FUNCTION: This peptide both inhibits the activity of the  
CC angiotensin-converting enzyme and enhances the action of  
CC bradykinin by inhibiting the kinases that inactivate it.  
CC It acts as an indirect hypotensive agent.  
DR PIR; G37196; G37196.  
KW Hypotensive agent; Pyrrolidone carboxylic acid.  
FT MOD RES 1  
SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;  
  
Query Match 18.4%; Score 7; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 P 3  
Db 5 P 5  
  
RESULT 14  
EI04 LITRU STANDARD; PRT; 5 AA.  
AC P82100;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Electrin 4.  
OS Litoria rubella (Desert tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;  
OC Pelodyadinae; Litoria.  
OX NCBI\_TaxID=104895;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Skin secretion;  
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
RT "Peptides from the skin glands of the Australian buzzing tree frog  
RT Litoria electrica. Comparison with the skin peptides from Litoria  
RT rubella.";  
RL Aust. J. Chem. 52:639-645(1999).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Skin.

Sun Sep 5 11:41:05 2004

KW Amphibian defense peptide; Amidation.  
 FT MOD RES 5 5 AMIDATION.  
 SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;  
 Query Match 18.4%; Score 7; DB 1; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LT 5  
 Db 2 IT 3  
 RESULT 15  
 PAP2 PARMA STANDARD; PRT; 5 AA.  
 AC P81864;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Pardaxin II (PXII) (Fragment).  
 OS Pardachirus marmoratus (Red sea Moses sole).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 OC Soleiidae; Soleidae; Pardachirus.  
 OX NCBI\_TaxID=31087;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin secretion;  
 RX MEDLINE=87057369; PubMed=3782138;  
 RA Lazarovici P., Primor N., Loew L.M.;  
 RT "Purification and pore-forming activity of two hydrophobic  
 RT polypeptides from the secretion of the Red sea Moses sole (Pardachirus  
 RT marmoratus).";  
 RL J. Biol. Chem. 261:16704-16713(1986).  
 CC -1- FUNCTION: Exhibits unusual shark repellent and surfactant  
 CC properties. Forms voltage-dependent, ion-permeable channels  
 CC in membranes. At high concentration causes cell membrane lysis.  
 CC -1- SUBUNIT: Monomer. In aqueous solution exists as a tetramer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the pardaxin family.  
 KW Toxin.  
 FT NON\_TER 5 5  
 SQ SEQUENCE 5 AA; 614 MW; 7769C9C8100000 CRC64;  
 Query Match 18.4%; Score 7; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 P 3  
 Db 5 P 5  
 Search completed: September 5, 2004, 11:25:35  
 Job time : 16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:22:21; Search time 58 Seconds  
(without alignments)  
32.640 Million cell updates/sec

Title: US-09-761-636A-12  
Perfect score: 38  
Sequence: 1 CVPLTC 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 11

Minimum DB seq length: 0  
Maximum DB seq length: 6

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rhodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	31.6	5	13 P83308	P83308 gallus gall
2	9	23.7	6	5 P83569	P83569 sepiia offic
3	5	13.2	5	2 P83073	P83073 bacillus ce
4	5	13.2	6	2 P83533	P83533 lactobacill
5	5	13.2	6	10 P82181	P82181 spinacia ol
6	5	13.2	6	10 P82541	P82541 spinacia ol
7	5	13.2	6	10 P82182	P82182 spinacia ol
8	4	10.5	4	5 P83568	P83568 sepiia offic
9	4	10.5	4	11 Q08433	Q08433 rattus sp.
10	2	5.3	5	10 Q99007	Q99007 hordeum vul
11	0	0.0	2	5 P83570	P83570 sepiia offic

## ALIGNMENTS

RESULT 1  
P83308 PRELIMINARY; PRT; 5 AA.  
AC P83308;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE FMRamide-like neuropeptide (LPLRF-amide).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]\_TaxID=9031;  
RS SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Brain;  
RX PubMed=613771;  
RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;  
RT "A novel active pentapeptide from chicken brain identified by  
RT antibodies to FMRamide."  
RL Nature 305:328-330(1983).  
CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)  
CC FAMILY.  
DR GO: GO:0007218; P-neuropeptide signaling pathway; TAS.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 5 5  
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 31.6%; Score 12; DB 13; Length 5;  
Best Local Similarity 66.7%; Pred. No. 1e+06;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VPL 4  
Db 1 LPL 3

RESULT 2  
P83569 PRELIMINARY; PRT; 6 AA.  
AC P83569;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Sperm attracting peptide Sepsap.  
OS Sepia officinalis (Common cuttlefish).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
OC Decapodiformes; Sepioidea; Sepiidae; Sepia.  
OX NCBI\_TaxID=6610;  
RN [1]  
RS SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND  
RP AMIDATION.  
RC TISSUE=Egg;  
RX PubMed=12207899;  
RA Zatylny C., Marvin L., Gagnon J., Henry J.;  
RT "Fertilization in Sepia officinalis: the first mollusk sperm-  
RT attracting peptide."  
RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).  
CC -!- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE  
CC COLLISION.  
CC -!- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES DURING  
CC VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED  
CC DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.  
CC ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.  
CC -!- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.  
KW Amidation.  
FT MOD\_RES 6 6  
SQ SEQUENCE 6 AA; 597 MW; 72C8676AA0470000 CRC64;

Query Match 23.7%; Score 9; DB 5; Length 6;  
Best Local Similarity 50.0%; Pred. No. 1e+06;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PL 4  
1 PI 2  
Db

RESULT 3  
P83073 PRELIMINARY; PRT; 5 AA.  
AC P83073;  
DT 01-OCT-2001 (TREMELrel. 18, Created)  
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE 88 kDa protein (Fragment).  
OS Bacillus cereus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1396;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=NCIMB 11796;  
RA Browne N., Dowds B.C.A.;  
RL Submitted (JUL-2001) to Swiss-Prot.  
FT NON TER 5  
SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;  
Query Match 13.2%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 T 5  
4 T 4  
Db

RESULT 4  
P83533 PRELIMINARY; PRT; 6 AA.  
AC P83533;  
DT 01-JUN-2003 (TREMELrel. 24, Created)  
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Unknown protein from 2D-page (Fragment).  
OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OX NCBI\_TaxID=1625;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=DSM 20451;  
RX PubMed=12112860;  
RA Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;  
RT "High pressure effects step-wise altered protein expression in Lactobacillus sanfranciscensis";  
RL Proteomics 2:765-774(2002).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN PROTEIN IS: 15 KDA.  
FT NON TER 1  
FT NON TER 1  
SQ SEQUENCE 6 AA; 590 MW; 6DDDD452D1AAC000 CRC64;  
Query Match 13.2%; Score 5; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 T 5  
2 T 2  
Db

RESULT 5  
P82181 PRELIMINARY; PRT; 6 AA.  
ID P82181  
AC P82181;

DT 01-JUN-2000 (TREMELrel. 14, Created)  
DT 01-JUN-2000 (TREMELrel. 14, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Amaranthaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0019843; F:rRNA binding; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR InterPro; IPR002363; Ribosomal L10eub.  
DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.  
KW Ribosomal protein, Chloroplast; rRNA-binding.  
FT NON TER 6  
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;  
Query Match 13.2%; Score 5; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 T 5  
5 T 5  
Db

RESULT 6  
P82541 PRELIMINARY; PRT; 6 AA.  
ID P82541;  
AC P82541;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Amaranthaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
RX MEDLINE=20435797; PubMed=10874039;  
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in the small subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 37:28455-28465(2000).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.  
CC -1- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.  
CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA FORM IS THE MINOR BASIC FORM.  
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.  
CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0019843; F:rRNA binding; IEA.



DR GO: 0003735; F: structural constituent of ribosome; IEA.  
 DR InterPro: IPR002222; Ribosomal S19.  
 DR PROSITE: PS00323; RIBOSOMAL\_S19; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON TER 6  
 SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 13.2%; Score 5; DB 10; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 T 5  
 Db 1 T 1

RESULT 7  
 ID P82182 PRELIMINARY; PRT; 6 AA.  
 AC P82182;  
 DT 01-JUN-2000 (TRENBLrel. 14, Created)  
 DT 01-JUN-2000 (TRENBLrel. 14, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
 OS Spinacia oleracea (Spinach).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Caryophyllales; Amaranthaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=cv. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 RT the 50 S subunit of an organelle ribosome (Chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR GO: 0009507; C: chloroplast; IEA.  
 DR GO: 0019843; F: rRNA binding; IEA.  
 DR GO: 0003735; F: structural constituent of ribosome; IEA.  
 DR InterPro: IPR002363; Ribosomal L10eub.  
 DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON TER 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 13.2%; Score 5; DB 10; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 T 5  
 Db 5 T 5

RESULT 8  
 ID P83568 PRELIMINARY; PRT; 4 AA.  
 AC P83568;  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Pheromone peptide ILME.  
 OS Sepia officinalis (Common cuttlefish).  
 CC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
 CC Decapodiformes; Sepioidae; Sepiidae; Sepia.  
 OX NCBI\_TaxID=6610;  
 RN [1]

RP SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS SPECTROMETRY.  
 RP TISSUE=Egg;  
 RX PubMed=10944467;  
 RA Zatylny C., Gagnon J., Boucaud-Camou E., Henry J.;  
 RT "ILME: a waterborne pheromonal peptide released by the eggs of Sepia officinalis";  
 RL Biochem. Biophys. Res. Commun. 275:217-222(2000).  
 RN [2]  
 RP SEQUENCE.

RC TISSUE=Egg;  
 RX PubMed=12207899;  
 RA Zatylny C., Marvin L., Gagnon J., Henry J.;  
 RT "Fertilization in Sepia officinalis: the first mollusk sperm-attracting peptide";  
 RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).  
 CC -!- FUNCTION: HAS MYOTROPIC ACTIVITY TARGETING THE GENITAL TRACT.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- TISSUE SPECIFICITY: FOLLICLE, FULLY GROWN OOCYTE AND EGG(EC2).  
 CC -!- MASS SPECTROMETRY: MW=505.4; METHOD=MALDI.  
 DR GO: 0005186; F: pheromone activity; IEA.  
 KW Pheromone.  
 SQ SEQUENCE 4 AA; 505 MW; 6B169720300000000 CRC64;

Query Match 10.5%; Score 4; DB 5; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 L 4  
 Db 2 L 2

RESULT 9  
 ID Q08433 PRELIMINARY; PRT; 4 AA.  
 AC Q08433;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Bilirubin UDP-glucuronosyltransferase (Fragment).  
 OS Rattus sp.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Gunn;  
 RX MEDLINE=91282758; PubMed=1840486; Koizumi O.;  
 RA Sato H., Aono S., Kashiwamata S., Koizumi O.;  
 RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the hyperbilirubinemic Gunn rat";  
 RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
 DR EMBL: S38636; AAB19259.1;  
 DR GO: 0016740; F: transferase activity; IEA.  
 KW Transferase.  
 FT NON TER 1  
 SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 10.5%; Score 4; DB 11; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 V 2  
 Db 2 V 2

RESULT 10  
 ID Q99007 PRELIMINARY; PRT; 5 AA.  
 AC Q99007;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)

Sun Sep 5 11:41:05 2004

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Alpha amylase (Fragment).  
 GN AMY1 GENE.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=91329704; PubMed=1831055;  
 RA Jacobsen J.V., Close T.J.;  
 RT "Control of transient expression of chimeric genes by gibberellic  
 RT acid and abscisic acid in protoplasts prepared from mature barley  
 RT aleurone layers.";  
 RL Plant Mol. Biol. 16:713-721(1991).  
 DR EMBL; X54643; CAA38455.1; -.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
 Query Match 5.3%; Score 2; DB 10; Length 5;  
 Best Local Similarity 0.0%; Pred. No. 1e+06;  
 Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 L 4  
 Db 1 M 1  
 RESULT 11  
 ID P83570 PRELIMINARY; PRT; 2 AA.  
 AC P83570;  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Neuropeptide Gwa.  
 OS Sepia officinalis (Common cuttlefish).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
 OC Decapodiformes; Sepioidea; Sepiidae; Sepia.  
 OX NCBI\_TaxID=6610;  
 RN [1]  
 RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.  
 RC TISSUE=Optic lobe;  
 RX PubMed=9437704;  
 RA Henry J., Favrel P., Boucaud-Camou E.;  
 RT "Isolation and identification of a novel Ala-pro-Gly-Tyr-amide-related  
 RT peptide inhibiting the motility of the mature oviduct in the  
 RT cuttlefish, Sepia officinalis.";  
 RL Peptides 18:1469-1474(1997).  
 CC -|- FUNCTION: REGULATORY NEUROPEPTIDE WITH MYOTROPIC ACTIVITY  
 CC TARGETING THE DISTAL OVIDUCT. INHIBITS THE MOTILITY OF THE OVIDUCT  
 CC BY DECREASING TONUS, FREQUENCY AND AMPLITUDE OF CONTRACTIONS.  
 CC -|- MASS SPECTROMETRY: MW=259.9; METHOD=MALDI.  
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 2 2  
 SQ SEQUENCE 2 AA; 261 MW; 7378100000000000 CRC64;  
 Query Match 0.0%; Score 0; DB 5; Length 2;  
 Best Local Similarity 0.0%; Pred. No. 1e+06;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 P 3  
 Db 1 G 1

Search completed: September 5, 2004, 11:26:39  
 Job time : 58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:19:05 ; Search time 62 seconds  
(without alignments)  
27.343 Million cell updates/sec

Title: US-09-761-636A-12

Perfect score: 38

Sequence: 1 CVP LTC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 59163

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0			AAU04531	VEGF base
2	26	68.4	6	4	AAU04531	VEGF base
3	25	65.8	6	3	AAU04531	VEGF base
4	25	65.8	6	3	AAU04531	VEGF base
5	25	65.8	6	3	AAU04531	VEGF base
6	24	63.2	6	3	AAU04531	VEGF base
7	23	60.5	6	3	AAU04531	VEGF base
8	23	60.5	6	3	AAU04531	VEGF base
9	23	60.5	6	3	AAU04531	VEGF base
10	23	60.5	6	3	AAU04531	VEGF base
11	22	57.9	4	2	AAU04531	VEGF base
12	22	57.9	4	2	AAU04531	VEGF base
13	22	57.9	4	2	AAU04531	VEGF base
14	22	57.9	4	2	AAU04531	VEGF base
15	22	57.9	4	2	AAU04531	VEGF base
16	22	57.9	4	2	AAU04531	VEGF base
17	22	57.9	4	2	AAU04531	VEGF base
18	22	57.9	4	2	AAU04531	VEGF base
19	21	55.3	6	3	AAU04531	VEGF base
20	21	55.3	6	3	AAU04531	VEGF base
21	20	52.6	6	3	AAU04531	VEGF base
22	20	52.6	6	3	AAU04531	VEGF base
23	20	52.6	6	3	AAU04531	VEGF base
24	20	52.6	6	3	AAU04531	VEGF base
25	20	52.6	6	3	AAU04531	VEGF base

26	20	52.6	5	3	AAU04531	VEGF base
27	20	52.6	5	4	AAU04531	VEGF base
28	20	52.6	5	5	AAU04531	VEGF base
29	20	52.6	5	5	AAU04531	VEGF base
30	20	52.6	5	5	AAU04531	VEGF base
31	20	52.6	5	5	AAU04531	VEGF base
32	20	52.6	6	2	AAU04531	VEGF base
33	20	52.6	6	2	AAU04531	VEGF base
34	20	52.6	6	2	AAU04531	VEGF base
35	20	52.6	6	2	AAU04531	VEGF base
36	20	52.6	6	2	AAU04531	VEGF base
37	20	52.6	6	2	AAU04531	VEGF base
38	20	52.6	6	2	AAU04531	VEGF base
39	20	52.6	6	2	AAU04531	VEGF base
40	20	52.6	6	3	AAU04531	VEGF base
41	20	52.6	6	3	AAU04531	VEGF base
42	20	52.6	6	3	AAU04531	VEGF base
43	20	52.6	6	5	AAU04531	VEGF base
44	19	50.0	4	2	AAU04531	VEGF base
45	19	50.0	5	2	AAU04531	VEGF base

## ALIGNMENTS

RESULT 1

AAU04531

ID AAU04531 standard; peptide; 6 AA.

AC AAU04531;

DT 26-SEP-2001 (first entry)

XX VEGF based monocyclic peptide 9.

XX Human; VEGF; vascular endothelial growth factor; angiogenesis;

KW neovascularisation; lymphangiogenesis; psoriasis; tumour;

KW diabetes induced neovascular sequelae; rheumatoid arthritis;

KW diabetic retinopathy; chronic inflammation; cyclic.

XX Synthetic.

XX Key

PH Disulfide-bond 1..6

FT Location/Qualifiers

FT /note= "This bond cyclises the peptide"

XX WO200152875-A1.

PD 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US001533.

PF 18-JAN-2001; 2000US-0176293P.

PR 16-MAY-2000; 2000US-0204590P.

XX (LUDW-) LUDWIG INST CANCER RES.

PA Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with angiogenesis,

PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment

PT from an exposed loop of a growth factor protein by oxidizing the cysteine

PT residues.

XX Claim 49; Page 32; 102pp; English.

PS The sequence represents a monomeric monocyclic peptide of the invention,

XX whose 3-dimensional structure is modelled on the exposed loop of human

CC VEGF (vascular endothelial growth factor). The invention relates to a

CC method of producing a monomeric monocyclic peptide by a measuring beta-

CC beta carbon separation distances on opposite antiparallel strands of a

peptide loop fragment from an exposed loop of a growth factor protein and cyclising the peptide by oxidising the cysteine residues. The monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic peptide with at least one amino acid deleted prior to cyclisation are used to interfere with angiogenesis, neovascularisation or lymphangiogenesis in a mammal with a condition characterised by angiogenesis, neovascularisation or lymphangiogenesis. The condition is diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised malignant or benign tumour, post-recovery cerebrovascular accident, post-angioplasty restenosis, head, heat or cold trauma, substance-induced neovascularisation of the liver, excessive hormone-related angiogenic dysfunction, diabetes induced neovascular sequelae, hypertension induced neovascular sequelae, or chronic liver infection. The peptides are also used to modulate vascular permeability in a mammal (the mammal has a condition characterised by fluid accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura, or brain. The peptides are used to image blood vessels and lymphatic vasculature. The monomeric and bicyclic peptides are used to interfere with at least one biological activity induced by VEGF, VEGF-C or -D and are also used in combination with an anti-inflammatory agent, to treat a chronic inflammation, especially rheumatoid arthritis, psoriasis and diabetic retinopathy

Sequence 6 AA;  
Query Match 100.0%; Score 38; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | | |  
Db 1 CVPLTC 6

RESULT 2  
AA561488  
ID AAY61488 standard; peptide; 6 AA.  
XX  
AC AAY61488;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1374.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.

XX Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..6  
XX  
PN WO957149-A2.  
XX  
PD 11-NOV-1999.  
XX  
PF 05-MAY-1999; 99WO-CA000363.  
XX  
PR 05-MAY-1998; 98US-00073040.  
PR 06-NOV-1998; 98US-00187859.  
PR 20-JAN-1999; 99US-00234395.  
PR 08-MAR-1999; 99US-00264516.  
XX  
PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
PI Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.  
XX New cadherin modulating agents, used for modulating nonclassical cadherin-mediated functions for treating e.g. cancers, obesity, rheumatoid arthritis, multiple sclerosis, diabetes or a neurological disease.  
XX Claim 36; Page 172; 252pp; English.  
XX The present invention describes cadherin modulating agents (MA) comprising peptides which comprise a nonclassical cadherin cell adhesion recognition (CAR) sequence. The MAs can be used for modulating nonclassical cadherin-mediated functions. They can be used for e.g. inhibiting adhesion of nonclassical-cadherin expressing cells in a mammal, enhancing delivery of a drug to a tumour in a mammal, treating cancer in a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-expressing cell, preventing or treating obesity in a mammal, stimulating blood vessel regression in a mammal, enhancing drug delivery to the central nervous system, treating a demyelinating neurological disease, increasing vasopermeability in a mammal, enhancing adhesion of nonclassical cadherin-expressing cells, inhibiting synaptic stability in a mammal, or preventing pregnancy in a mammal. They can also be used for e.g. enhancing or directing neurite outgrowth, facilitating wound healing or reducing scar tissue, or enhancing adhesion of foreign tissue in a mammal. They can also be used for treating e.g. psoriasis, arthritis, age-related macular degeneration, multiple sclerosis and diabetes. The products can also be used for detection and diagnosis and in bioreactors. AAY60592 to AAY64572 represent specifically claimed peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in the exemplification of the present invention

Sequence 6 AA;

Query Match 68.4%; Score 26; DB 3; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | | |  
Db 1 CEPKTC 6

RESULT 3  
AAY62223  
ID AAY62223 standard; peptide; 6 AA.  
XX  
AC AAY62223;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:3970.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.  
XX Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..6  
XX  
PN WO957149-A2.  
XX  
PD 11-NOV-1999.  
XX

PF 05-MAY-1999; 99WO-CA000363.  
 XX  
 PR 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 DR  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX  
 PS Claim 54; Page 184; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 SQ Sequence 6 AA;  
 Query Match 65.8%; Score 25; DB 3; Length 6;  
 Best Local Similarity 66.7%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CVPLTC 6  
 DB 1 CDPKTC 6  
 RESULT 4  
 AAY62763  
 ID AAY62763 standard; peptide; 6 AA.  
 XX  
 AC AAY62763;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:4046.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..6  
 XX  
 PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 05-MAY-1999; 99WO-CA000363.  
 XX  
 PR 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 DR  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX  
 PS Claim 72; Page 193; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 SQ Sequence 6 AA;  
 Query Match 65.8%; Score 25; DB 3; Length 6;  
 Best Local Similarity 66.7%; Pred. NO. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CVPLTC 6  
 DB 1 CDPKTC 6  
 RESULT 5  
 AAY62006  
 ID AAY62006 standard; peptide; 6 AA.  
 XX  
 AC AAY62006;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-12 cell adhesion recognition cyclic peptide SEQ ID NO:1799.



SQ Sequence 6 AA;

Query Match 63.2%; Score 24; DB 3; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
 | : | | |  
 Db 1 CIVLAC 6

RESULT 7  
 AAY85373  
 ID AAY85373 standard; peptide; 6 AA.  
 XX  
 AC AAY85373;

DT 19-JUN-2000 (first entry)

DE IL-2 derived anti-inflammatory peptide pep11.

KW Interleukin-2; IL-2; antiinflammatory; antiarthritic; antirheumatic;  
 KW antidiabetic; neuroprotective; dermatological; immunosuppressive;  
 KW optalmological; autoimmune disease; multiple sclerosis; uveitis;  
 KW systemic lupus erythematosus; Crohn's disease.

OS Synthetic.  
 OS Homo sapiens.

PN WO200011028-A2.

PD 02-MAR-2000.

PF 19-AUG-1999; 99WO-IL000448.

PR 21-AUG-1998; 98GB-00018370.

PR 31-AUG-1998; 98IL-00126009.

PR 16-MAY-1999; 99IL-00129980.

PA (YEDA ) YEDA RES & DEV CO LTD.

PI Lider O, Ariel A, Hershkoviz R, Yavin EJ, Fridkin M;

DR WPI; 2000-256367/22.

PT Synthetic antiinflammatory peptide derived from IL-2 and its derivatives  
 useful for treating inflammatory autoimmune diseases such as rheumatoid  
 arthritis, multiple sclerosis and systemic lupus erythematosus.

PS Claim 6; Page 35; 49pp; English.

XX The invention provides synthetic antiinflammatory peptides derived from  
 interleukin-2 (IL-2). They can be used for inhibition of adhesion of  
 activated T-cells to ECM proteins such as fibronectin, laminin, collagen  
 type-IV; inhibition of chemotactic migration of T-cell through ECM  
 proteins preferably fibronectin; inhibition of cytokine or mitogen  
 induced T-cell proliferation; inhibition of spontaneous or induced,  
 preferably TNF-alpha induced cytokine secretion (e.g. IL-8, IL-1beta) by  
 stimulated T-cells and intestinal epithelial cells. The anti-inflammatory  
 peptides and their derivatives are useful for preparing compositions for  
 treating and/or alleviating chronic or acute inflammatory disorders and  
 autoimmune diseases such as rheumatoid arthritis, diabetes type-I,  
 multiple sclerosis, systemic lupus erythematosus, bowel inflammation,  
 uveitis, and Crohn's disease. Sequences AAY85366-374 represent modified  
 anti-inflammatory derivative peptides derived from the IL-2 derived anti-  
 inflammatory peptide pep1 (AAY85363)

XX Sequence 6 AA;

Query Match 60.5%; Score 23; DB 3; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
 | : | | |  
 Db 1 CIVLAC 6

RESULT 8  
 AAY62757  
 ID AAY62757 standard; peptide; 6 AA.  
 XX  
 AC AAY62757;

DT 02-MAR-2000 (first entry)

DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:4040.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

OS Synthetic.  
 OS Homo sapiens.

FN Key Location/Qualifiers  
 FT Disulfide-bond 1..6

PN WO9957149-A2.

PD 11-NOV-1999.

PF 05-MAY-1999; 99WO-CA000363.

PR 05-MAY-1998; 98US-00073040.

PR 06-NOV-1998; 98US-00187859.

PR 20-JAN-1999; 99US-00234395.

PR 08-MAR-1999; 99US-00264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

DR WPI; 2000-038791/03.

PT New cadherin modulating agents, used for modulating nonclassical cadherin  
 mediated functions for treating e.g. cancers, obesity, rheumatoid  
 arthritis, multiple sclerosis, diabetes or a neurological disease.

PS Claim 72; Page 193; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)  
 comprising peptides which comprise a nonclassical cadherin cell adhesion  
 recognition (CAR) sequence. The MAs can be used for modulating  
 nonclassical cadherin-mediated functions. They can be used for e.g.  
 inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 mammal, enhancing delivery of a drug through the skin of a mammal,  
 enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 expressing cell, preventing or treating obesity in a mammal, stimulating  
 blood vessel regression in a mammal, enhancing drug delivery to the  
 central nervous system, treating a demyelinating neurological disease,  
 increasing vasopermeability in a mammal, enhancing adhesion of  
 nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 a mammal, or preventing pregnancy in a mammal. They can also be used for  
 e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 -related macular degeneration, multiple sclerosis and diabetes. The  
 products can also be used for detection and diagnosis and in bioreactors.

CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in  
 CC the exemplification of the present invention

XX SQ Sequence 6 AA;

Query Match 60.5%; Score 23; DB 3; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
 | | | |  
 Db 1 CDELTC 6

RESULT 9

AAV62488  
 ID AAY62488 standard; peptide; 6 AA.

XX AC AAY62488;

XX DT 02-MAR-2000 (first entry)

XX DE Cadherin-15 cell adhesion recognition cyclic peptide SEQ ID NO:2154.

XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;

XX KW Inhibition; cadherin extracellular domain; cell adhesion recognition;  
 XX KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 XX KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;

XX KW cadherin related neuronal receptor; LI-cadherin; protocadherin;

XX KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

XX KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;

XX KW neurological disease; cyclic.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Disulfide-bond 1..6

XX WO9957149-A2.

XX PN 11-NOV-1999.

XX PD 05-MAY-1999; 99WO-CA000363.

XX PF 05-MAY-1998; 98US-00073040.

XX PR 06-NOV-1998; 98US-00187859.

XX PR 20-JAN-1999; 99US-00234395.

XX PR 08-MAR-1999; 99US-00264516.

XX XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX XX Blaschuk OW, Gour BJ, Byers S;

XX XX WPI; 2000-038791/03.

XX XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT-mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.

XX PS Claim 60; Page 188; 252pp; English.

XX CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating

CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioassays.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in  
 CC the exemplification of the present invention

XX SQ Sequence 6 AA;

Query Match 60.5%; Score 23; DB 3; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
 | | | |  
 Db 1 CDELTC 6

RESULT 10

AAV44431

ID AAY44431 standard; peptide; 6 AA.

XX AC AAY44431;

XX DT 22-MAR-2000 (first entry)

XX DE Peptide 1 derived from domain 1 of human beta-2 glycoprotein I.

XX KW Human beta-2 glycoprotein I; beta-2 GPI; toleragen; B cell anergy;

XX KW beta-2 GPI-dependent antiphospholipid antibody; thrombosis;

XX KW recurrent foetal loss; thrombocytopenia; autoimmune disease;

XX KW systemic lupus erythematosus; coagulation assay.

XX OS Homo sapiens.

XX PN WO9964595-A1.

XX PD 16-DEC-1999.

XX PF 09-JUN-1999; 99WO-US013194.

XX PR 09-JUN-1998; 98US-0088656P.

XX PR 05-OCT-1998; 98US-0103088P.

XX PR 08-JUN-1999; 99US-00328199.

XX XX (LJOL-) LA JOLLA PHARM CO.

XX XX Marquis DM, Iverson GM, Victoria BJ, Jones DS, Linnik MD;

XX XX WPI; 2000-116542/10.

XX XX New isolated domain 1 beta-2 GPI polypeptides, used for inhibiting  
 PT antiphospholipid antibodies for treating, e.g. thrombosis.

XX PS Claim 3; Page 19; 158pp; English.

XX CC The present sequence is a peptide fragment derived from domain 1 of human  
 CC beta-2 glycoprotein, a phospholipid binding serum protein. Isolated  
 CC domain 1 of beta-2 GPI protein binds to and inhibits beta-2 GPI-dependent  
 CC antiphospholipid antibodies. The fragments are useful as toleragens when  
 CC they bind to the antibodies at the surface of a B cell and triggers B  
 CC cell anergy. The polypeptides and mimetics can be used for treating  
 CC disorders associated with beta 2GPI-dependent API-associated pathologies,  
 CC e.g. thrombosis, recurrent foetal loss, thrombocytopenia or autoimmune  
 CC diseases such as systemic lupus erythematosus. The polypeptides can also  
 CC be used to detect and purify antibodies. They can also be used in



```

CC coagulation assays
XX
SQ Sequence 6 AA;

Query Match      60.5%; Score 23; DB 3; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 CTRVC 6

RESULT 11
AAR15772
ID AAR15772 standard; protein; 4 AA.
XX
AC AAR15772;
XX
DT 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 29-JAN-1992 (first entry)
XX
DE Farnesyl-protein transferase inhibitor (25).
XX
KW Farnesyl; transferase; FT; inhibitor; p21ras; rat.
XX
OS Synthetic.
XX
PN WO9116340-A.
XX
PD 31-OCT-1991.
XX
PF 18-APR-1990; 90US-00510706.
XX
PR 18-APR-1990; 90US-00510706.
PR 20-NOV-1990; 90US-00615715.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Brown MS, Goldstein JL, Reiss Y;
XX
DR WPI; 1991-339750/46.
XX
CC Composn. comprising purified farnesyl-protein transferase - used to
inhibit attachment of farnesyl moiety to RAS protein in malignant cells
and to treat cancer.
XX
PS Claim 25; Page 68; 87pp; English.
XX
CC This peptide or the peptides represented in AAR15751-81, AAR14723 and
AAR14711 inhibit the rat FTs represented in AAR14712-22. They show FT
inhibition at an IC50 of 0.01-10 microm. The most potent inhibitors are
ones in which phenylalanine occurs at the third position of a
tetrapeptide whose N-terminus is cysteine. The inhibitors have a farnesyl
acceptor or inhibitor sequence within its structure and are capable of
inhibiting the farnesylation of p21ras by FT. See also AAR14711-23 and
AAQ14541-47. (Updated on 09-JAN-2003 to add missing OS field.) (Updated
on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 4 AA;

Query Match      57.9%; Score 22; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPL 4
Db 1 CVPM 4

RESULT 12
AAR49769
ID AAR49769 standard; peptide; 4 AA.
XX
AC AAR49769;
XX
DT 25-MAR-2003 (revised)
DT 08-AUG-1994 (first entry)
XX
DE Farnesyltransferase-inhibitor.
XX
KW Farnesyltransferase-inhibitor; farnesyltransferase; FT; p21ras;
ras protein; farnesylation; cancer therapy.
XX
OS Synthetic.
XX
PN WO9404561-A1.
XX
PD 03-MAR-1994.
XX
PF 24-AUG-1993; 93WO-US008062.
XX
PR 24-AUG-1992; 92US-00935087.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (GETH ) GENENTECH INC.
XX
PI Brown MS, Goldstein JL, Reiss Y, Marsters JC;
XX
DR WPI; 1994-083105/10.
XX
CC New farnesyl-transferase inhibitors - used for inhibiting attachment of a
farnesyl moiety to a p21ras protein in malignant cells.
XX
PS Disclosure; Page 33; 183pp; English.
XX
CC Peptides given in AAR49741-75, AAR49777-78 and AAR49785-88, which include
a family of tetrapeptides based on the recognition site (AAR49776) of
farnesyltransferase (FT), are potential anticancer agents that inhibit
FT, thereby preventing expression of p21ras. (Updated on 25-MAR-2003 to
correct PN field.)
XX
SQ Sequence 4 AA;

Query Match      57.9%; Score 22; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPL 4
Db 1 CVPM 4

RESULT 13
AAR77833
ID AAR77833 standard; protein; 4 AA.
XX
AC AAR77833;
XX
DT 25-MAR-2003 (revised)
DT 23-JAN-1996 (first entry)
XX
DE Farnesyl transferase inhibitor tetrapeptide, CVPM.
XX
KW Farnesyl transferase; inhibitor; cancer; ras; p21.
XX
OS Synthetic.
XX
PN US5420245-A.
XX
PD 30-MAY-1995.
XX
PF 03-APR-1992; 92US-00863169.
XX
PR 18-APR-1990; 90US-00510706.

```

PR 20-NOV-1990; 90US-00615715.  
 PR 16-JAN-1992; 92US-00822011.  
 PA (TEXA ) UNIV TEXAS.  
 XX Reiss Y, Goldstein JL, Brown MS;  
 PI WPI; 1995-206308/27.  
 DR  
 XX  
 XX New farnesyl transferase inhibitor peptide(s) - based on farnesyl  
 PT acceptor substrate carboxy terminal sequences, used for the treatment of  
 PT cancer.  
 XX  
 PS Claim 2; Col 62; 55pp; English.  
 XX  
 CC AAR77800, AAR77805-R77838 are tetrapeptide inhibitors of farnesyl  
 CC transferase. They all obey a generic formula for the C-terminal sequence  
 CC of 4-10 amino acid inhibitory peptides; the formula is -CAAX, where C=  
 CC cysteine, A= any aliphatic, aromatic or hydroxy amino acid and X= any  
 CC normal amino acid. Farnesyl transferase is involved in the farnesylation  
 CC of various cellular proteins including the cancer related ras proteins.  
 CC The transforming activity of ras is dependent on the localisation of the  
 CC protein to membranes, a property which is thought to be dependent upon  
 CC the addition of farnesyl groups. The peptide inhibitors are useful for  
 CC treating cancers and ras-related cancers in particular. (Updated on 25-  
 CC MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 4 AA;

Query Match 57.9%; Score 22; DB 2; Length 4;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPL 4  
 |||:  
 Db 1 CVPM 4

## RESULT 14

AAW04459  
 ID AAW04459 standard; peptide; 4 AA.

AC AAW04459;  
 XX  
 DT 30-JUL-1997 (first entry)  
 XX  
 DE Farnesyl transferase peptide inhibitor used in cancer treatment.  
 XX  
 KW Farnesyl transferase; inhibitor; cancer; tumour; neoplasia; prenyl;  
 KW ras protein; K-ras B; malignant; detection; identification.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9634113-A2.  
 XX  
 PD 31-OCT-1996.  
 XX  
 PF 29-APR-1996; 96WO-US005969.  
 XX  
 XX 27-APR-1995; 95US-00429964.  
 XX

(TEXA ) UNIV TEXAS SYSTEM.

PI Brown MS, Goldstein JL, James GL;

XX WPI; 1996-497642/49.

XX Assay for farnesyl transferase activity - by determining ability to  
 PT transfer farnesyl moiety to K-Ras B protein, partic. useful for  
 PT identifying inhibitors.

XX Disclosure; Page 34; 257pp; English.

XX

CC AAW04433-W04465 are peptide inhibitors of farnesyl transferase (FT)  
 CC activity. The peptides block the attachment of prenyl groups to ras  
 CC proteins in malignant cells of patients suffering from cancer or a  
 CC precancerous state and as such are used to treat cancer. The peptides  
 CC were identified by determining the ability of candidate substances to  
 CC inhibit a FT enzyme, by inhibiting the transfer of a farnesyl moiety to a  
 CC K-RasB protein  
 XX  
 SQ Sequence 4 AA;

Query Match 57.9%; Score 22; DB 2; Length 4;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPL 4  
 |||:  
 Db 1 CVPM 4

## RESULT 15

AAW67428  
 ID AAW67428 standard; peptide; 5 AA.

AC AAW67428;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE HCV peptide analogue #2.  
 XX  
 KW Hepatitis C virus; HCV; nucleocapsid; core protein; analogue; antibody;  
 KW non-structural protein; thioamide bond; peptide bond.  
 XX  
 OS Synthetic.  
 OS Hepatitis C virus.

Key Location/Qualifiers

FT Modified-site 1 /note= "side chain protected by butyl group"

FT Modified-site 2 /note= "side chain protected by tosyl group"

FT Disulfide-bond 5

FT /note= "side chain protected by benzoyl group and C-terminus protected by benzyl group"

XX JPI0226698-A.

XX 25-AUG-1998.

XX 19-FEB-1997; 97JP-00034702.

XX 19-FEB-1997; 97JP-00034702.

XX (KYOW ) KYOWA MEDEX KK.

XX WPI; 1998-515103/44.

XX Determination of antibody in sample - uses peptide analog absorbed or  
 PT chemically bound on carrier as antigen.

XX Example 1; Page 9; 13pp; Japanese.

XX This sequence represents an analogue peptide of the Hepatitis C virus  
 CC (HCV) nucleocapsid core protein. The invention relates to peptide  
 CC analogues derived from HCV proteins, e.g. AAW67417-W67426, which can be  
 CC used for the determination of anti-HCV antibodies in a sample. Preferably  
 CC the peptide analogues contain one or more thioamide peptide bonds where  
 CC at least one oxygen atom of the peptide bond is replaced by sulphur atom.  
 CC The peptide analogues can be adsorbed or chemically bound to a carrier

XX Sequence 5 AA;

Query Match 57.9%; Score 22; DB 2; Length 5;

Best Local Similarity 80.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLT 5  
| | |  
Db 1 CRPLT 5

Search completed: September 5, 2004, 11:25:15  
Job time : 64 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:26:47 ; Search time 67 Seconds  
(without alignments)  
28.215 Million cell updates/sec

Title: US-09-761-636A-12  
Perfect score: 38  
Sequence: 1 CVPLTC 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 28303

Minimum DB seq length: 0  
Maximum DB seq length: 6

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	6	9	US-09-761-636A-12
2	26	68.4	6	14	US-10-006-869-1374
3	26	68.4	6	15	US-10-395-032-1374
4	25	65.8	6	14	US-10-006-869-1798
5	25	65.8	6	14	US-10-006-869-3970
6	25	65.8	6	14	US-10-006-869-4046
7	25	65.8	6	15	US-10-395-032-1798
8	25	65.8	6	15	US-10-395-032-3970
9	25	65.8	6	15	US-10-395-032-4046
10	25	65.8	6	16	US-10-422-571-75
11	24	63.2	6	14	US-10-006-869-3573
12	24	63.2	6	15	US-10-395-032-3573
13	23	60.5	6	14	US-10-006-869-2154
14	23	60.5	6	14	US-10-006-869-4040
15	23	60.5	6	15	US-10-395-032-2154

16	23	60.5	6	15	US-10-395-032-4040	Sequence 4040, Ap
17	22	57.9	4	14	US-10-083-894-35	Sequence 35, Appl
18	22	57.9	6	11	US-09-943-944E-103	Sequence 103, App
19	22	57.9	6	14	US-10-006-869-1737	Sequence 1737, Ap
20	22	57.9	6	14	US-10-006-869-2719	Sequence 2719, Ap
21	22	57.9	6	15	US-10-395-032-1737	Sequence 1737, Ap
22	22	57.9	6	15	US-10-395-032-2719	Sequence 2719, Ap
23	21	55.3	6	9	US-09-911-838-184	Sequence 184, App
24	21	55.3	6	9	US-09-911-838-186	Sequence 186, App
25	21	55.3	6	10	US-09-792-286-222	Sequence 222, App
26	21	55.3	6	10	US-09-792-286-226	Sequence 226, App
27	21	55.3	6	10	US-09-792-286-278	Sequence 278, App
28	21	55.3	6	14	US-10-006-869-1885	Sequence 1885, Ap
29	21	55.3	6	14	US-10-006-869-2747	Sequence 2747, Ap
30	21	55.3	6	14	US-10-058-513-9	Sequence 9, Appl
31	21	55.3	6	15	US-10-395-032-1885	Sequence 1885, Ap
32	21	55.3	6	15	US-10-395-032-2747	Sequence 2747, Ap
33	20	52.6	5	10	US-09-753-139C-8	Sequence 8, Appl
34	20	52.6	5	12	US-10-416-797-16	Sequence 16, Appl
35	20	52.6	6	9	US-09-911-838-90	Sequence 90, Appl
36	20	52.6	6	9	US-09-911-838-92	Sequence 92, Appl
37	20	52.6	6	9	US-09-911-838-94	Sequence 94, Appl
38	20	52.6	6	9	US-09-911-838-96	Sequence 96, Appl
39	20	52.6	6	12	US-09-935-430-697	Sequence 697, App
40	20	52.6	6	14	US-10-006-869-1591	Sequence 1591, Ap
41	20	52.6	6	14	US-10-271-343-52	Sequence 52, Appl
42	20	52.6	6	14	US-10-436-826-50	Sequence 50, Appl
43	20	52.6	6	14	US-10-277-292-697	Sequence 697, App
44	20	52.6	6	15	US-10-280-340-697	Sequence 697, App
45	20	52.6	6	15	US-10-395-032-1591	Sequence 1591, Ap

ALIGNMENTS

RESULT 1  
US-09-761-636A-12  
; Sequence 12, Application US/09761636A  
; Patent No. US20020065218A1  
; GENERAL INFORMATION:  
; APPLICANT: ACHEN, Marc  
; APPLICANT: STACKER, Steven  
; APPLICANT: HUGHES, Richard  
; APPLICANT: GENDERON, Angela  
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR  
; FILE REFERENCE: 1064/48505 Achen et al  
; CURRENT APPLICATION NUMBER: US/09/761,636A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/176,293  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/204,590  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 12  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-636A-12

Query Match 100.0%; Score 38; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
|||  
Db 1 CVPLTC 6

RESULT 2  
US-10-006-869-1374  
; Sequence 1374, Application US/10006869  
; Publication No. US20030082166A1

```
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1374
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-10-006-869-1374

Query Match      68.4%; Score 26; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CVELTC 6
Db 1 CEPKTC 6

RESULT 3
US-10-395-032-1374
; Sequence 1374, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1374
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-10-395-032-1374

Query Match      68.4%; Score 26; DB 15; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CVELTC 6
Db 1 CEPKTC 6

RESULT 4
US-10-006-869-1798
; Sequence 1798, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
```

```
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1798
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence
US-10-006-869-1798

Query Match      65.8%; Score 25; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CVELTC 6
Db 1 CDEKTC 6

RESULT 5
US-10-006-869-3970
; Sequence 3970, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3970
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence
US-10-006-869-3970

Query Match      65.8%; Score 25; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CVELTC 6
Db 1 CDEKTC 6

RESULT 6
US-10-006-869-4046
; Sequence 4046, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4046
; LENGTH: 6
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-10-006-869-4046

Query Match          65.8%; Score 25; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 CDPKTC 6

RESULT 7
US-10-395-032-1798
; Sequence 1798, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1798
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence
US-10-395-032-1798

Query Match          65.8%; Score 25; DB 15; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 CDPKTC 6

RESULT 8
US-10-395-032-3970
; Sequence 3970, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3970
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence
US-10-395-032-3970

Query Match          65.8%; Score 25; DB 15; Length 6;
```

```
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 CDPKTC 6

RESULT 9
US-10-395-032-4046
; Sequence 4046, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4046
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-10-395-032-4046

Query Match          65.8%; Score 25; DB 15; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 CDPKTC 6

RESULT 10
US-10-422-571-75
; Sequence 75, Application US/10422571
; Publication No. US20040083497A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20082.00
; CURRENT APPLICATION NUMBER: US/10/422,571
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US60/404,306
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US60/423,290
; PRIOR FILING DATE: 2002-11-01
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-422-571-75

Query Match          65.8%; Score 25; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 PLTC 6  
| | | |  
Db 3 PLTC 6

## RESULT 11

US-10-006-869-3573  
; Sequence 3573, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3573  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
; OTHER INFORMATION: recognition sequence  
US-10-006-869-3573

Query Match 63.2%; Score 24; DB 14; Length 6;  
Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPVSC 6

## RESULT 12

US-10-395-032-3573  
; Sequence 3573, Application US/10395032  
; Publication No. US20030229199A1  
; GENERAL INFORMATION:  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C9  
; CURRENT APPLICATION NUMBER: US/10/395,032  
; CURRENT FILING DATE: 2003-03-21  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3573  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
; OTHER INFORMATION: recognition sequence  
US-10-395-032-3573

Query Match 63.2%; Score 24; DB 15; Length 6;  
Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPVSC 6

## RESULT 13

US-10-006-869-2154  
; Sequence 2154, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2154  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-15 cell adhesion recognition sequence  
US-10-006-869-2154

Query Match 60.5%; Score 23; DB 14; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDELTC 6

## RESULT 14

US-10-006-869-4040  
; Sequence 4040, Application US/10006869  
; Publication No. US20030082166A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C7  
; CURRENT APPLICATION NUMBER: US/10/006,869  
; CURRENT FILING DATE: 2001-12-03  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4040  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-10-006-869-4040

Query Match 60.5%; Score 23; DB 14; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDELTC 6

## RESULT 15

US-10-395-032-2154  
; Sequence 2154, Application US/10395032  
; Publication No. US20030229199A1  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.



```
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2154
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-15 cell adhesion recognition sequence
US-10-395-032-2154

Query Match      60.5%; Score 23; DB 15; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CVPLTC 6
      | |||
Db      1 CDELTC 6

Search completed: September 5, 2004, 11:32:10
Job time : 68 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:25:22 ; Search time 20 Seconds  
(without alignments)  
15.488 Million cell updates/sec

Title: US-09-761-636A-12

Perfect score: 38

Sequence: 1 CVP LTC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 40703

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*

6: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	68.4	6	4	US-09-187-859-1374
2	26	68.4	6	4	US-09-839-542B-1374
3	25	65.8	6	4	US-09-187-859-1798
4	25	65.8	6	4	US-09-187-859-3970
5	25	65.8	6	4	US-09-187-859-4046
6	25	65.8	6	4	US-09-839-542B-1798
7	25	65.8	6	4	US-09-839-542B-3970
8	25	65.8	6	4	US-09-839-542B-4046
9	24	63.2	6	4	US-09-187-859-3573
10	24	63.2	6	4	US-09-839-542B-3573
11	23	60.5	6	4	US-09-187-859-2154
12	23	60.5	6	4	US-09-187-859-4040
13	23	60.5	6	4	US-09-839-542B-2154
14	23	60.5	6	4	US-09-839-542B-4040
15	22	57.9	4	2	US-08-429-964-41
16	22	57.9	4	4	US-09-665-362A-35
17	22	57.9	4	5	PCT-US93-08062-41
18	22	57.9	6	4	US-09-187-859-1737
19	22	57.9	6	4	US-09-187-859-2719
20	22	57.9	6	4	US-09-839-542B-1737
21	22	57.9	6	4	US-09-839-542B-2719
22	21	55.3	5	4	US-09-082-358B-90
23	21	55.3	5	4	US-09-050-861B-24
24	21	55.3	6	4	US-09-187-859-1885
25	21	55.3	6	4	US-09-187-859-2747
26	21	55.3	6	4	US-09-839-542B-1885
27	21	55.3	6	4	US-09-839-542B-2747

28 20 52.6 4 1 US-08-548-540-120 Sequence 120, App  
29 20 52.6 4 5 PCT-US96-09809-120 Sequence 120, App  
30 20 52.6 4 4 US-09-322-791-4 Sequence 4, Appl  
31 20 52.6 5 4 US-09-322-791-6 Sequence 6, Appl  
32 20 52.6 6 3 US-08-394-630-13 Sequence 13, Appl  
33 20 52.6 6 3 US-08-750-142B-21 Sequence 21, Appl  
34 20 52.6 6 4 US-09-187-859-1591 Sequence 1591, Ap  
35 20 52.6 6 4 US-09-839-542B-1591 Sequence 1591, Ap  
36 20 52.6 6 4 US-09-535-852-1344 Sequence 1344, Ap  
37 20 52.6 6 4 US-09-535-852-1670 Sequence 1670, Ap  
38 19 50.0 5 3 US-09-248-588-109 Sequence 109, App  
39 19 50.0 6 1 US-08-201-046A-4 Sequence 4, Appl  
40 19 50.0 6 1 US-08-433-318A-181 Sequence 181, App  
41 19 50.0 6 2 US-08-922-048-181 Sequence 181, App  
42 19 50.0 6 3 US-08-750-142B-23 Sequence 23, Appl  
43 19 50.0 6 4 US-09-460-384-20 Sequence 20, Appl  
44 19 50.0 6 4 US-09-187-859-738 Sequence 738, App  
45 19 50.0 6 4 US-09-187-859-824 Sequence 824, App

#### ALIGNMENTS

##### RESULT 1

US-09-187-859-1374  
; Sequence 1374, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1374  
; TYPE: PRT  
; LENGTH: 6  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence  
US-09-187-859-1374

Query Match 68.4%; Score 26; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 2;

QY 1 CVP LTC 6  
| | | |  
Db 1 CEPKTC 6

##### RESULT 2

US-09-839-542B-1374  
; Sequence 1374, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1374  
; LENGTH: 6  
; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence

US-09-839-542B-1374

Query Match 68.4%; Score 26; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

Qy 1 CVPLTC 6  
| | | |  
Db 1 CEPKTC 6

#### RESULT 3

US-09-187-859-1798  
; Sequence 1798, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

; FILE REFERENCE: 100086.407C1

; CURRENT APPLICATION NUMBER: US/09/187,859A

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 4052

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1798

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-12 cell adhesion recognition sequence

US-09-187-859-1798

Query Match 65.8%; Score 25; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

Qy 1 CVPLTC 6  
| | | |  
Db 1 CDPKTC 6

#### RESULT 4

US-09-187-859-3970  
; Sequence 3970, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

; FILE REFERENCE: 100086.407C1

; CURRENT APPLICATION NUMBER: US/09/187,859A

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 4052

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3970

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence

US-09-187-859-3970

Query Match 65.8%; Score 25; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

; Sequence 3970, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3970  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence  
US-09-839-542B-3970

Query Match 65.8%; Score 25; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPKTC 6

RESULT 8  
US-09-839-542B-4046  
; Sequence 4046, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4046  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542B-4046

Query Match 65.8%; Score 25; DB 4; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPKTC 6

RESULT 9  
US-09-187-859-3573  
; Sequence 3573, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2154

; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3573  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
; OTHER INFORMATION: recognition sequence  
US-09-187-859-3573

Query Match 63.2%; Score 24; DB 4; Length 6;  
Best Local Similarity 50.0%; Pred. No. 3e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPVSC 6

RESULT 10  
US-09-839-542B-3573  
; Sequence 3573, Application US/09839542B  
; Patent No. 6569996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3573  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion  
; OTHER INFORMATION: recognition sequence  
US-09-839-542B-3573

Query Match 63.2%; Score 24; DB 4; Length 6;  
Best Local Similarity 50.0%; Pred. No. 3e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6  
| | | |  
Db 1 CDPVSC 6

RESULT 11  
US-09-187-859-2154  
; Sequence 2154, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2154

```
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-15 cell adhesion recognition sequence
US-09-187-859-2154

Query Match          60.5%; Score 23; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPLTC 6
   | |||
Db 1 CDELTC 6

RESULT 12
US-09-187-859-4040
; Sequence 2154, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4040
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-15 cell adhesion recognition sequence
US-09-187-859-4040

Query Match          60.5%; Score 23; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPLTC 6
   | |||
Db 1 CDELTC 6

RESULT 13
US-09-839-542B-2154
; Sequence 2154, Application US/09839542B
; Patent No. 656996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2154
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-15 cell adhesion recognition sequence
US-09-839-542B-2154
```

```
Query Match          60.5%; Score 23; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPLTC 6
   | |||
Db 1 CDELTC 6

RESULT 14
US-09-839-542B-4040
; Sequence 4040, Application US/09839542B
; Patent No. 656996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4040
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-839-542B-4040

Query Match          60.5%; Score 23; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3e+05; 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVPLTC 6
   | |||
Db 1 CDELTC 6

RESULT 15
US-08-429-964-41
; Sequence 41, Application US/08429964
; Patent No. 5962243
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; APPLICANT: JAMES, GUY L.
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P. O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,964
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,625
; FILING DATE: 16-FEB-1993
```

```
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/822,011
; FILING DATE: ABANDONED
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/US/91/02650
; FILING DATE: 18-APR-1991
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/615,715
; FILING DATE: 20-NOV-1990
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/510,706
; FILING DATE: 18-APR-1990 (ABANDONED)
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:432/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-429-964-41

Query Match          57.9%; Score 22; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CVPL 4
Db      1 CVPM 4

Search completed: September 5, 2004, 11:30:16
Job time : 21 secs
```

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:31:03 ; Search time 23 Seconds

(without alignments)  
41.822 Million cell updates/sec

Title: US-09-761-636A-13

Perfect score: 56

Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 1101

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: Pirl:\*

2: Pirl:\*

3: Pirl:\*

4: Pirl:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	35.7	6	2	I65546
2	19	33.9	5	2	A60521
3	19	33.9	10	2	C39111
4	18	32.1	9	2	PT0247
5	17	30.4	10	2	A61622
6	17	30.4	10	2	PH0944
7	16	28.6	4	2	I51049
8	16	28.6	8	2	G33098
9	16	28.6	10	2	PH0926
10	16	28.6	10	2	A42689
11	15	26.8	8	2	A49712
12	15	26.8	9	2	S43630
13	15	26.8	10	2	E86128
14	15	26.8	10	2	B22565
15	14	25.0	5	2	A28340
16	14	25.0	7	2	I57018
17	14	25.0	8	2	A60356
18	14	25.0	9	2	S26508
19	14	25.0	9	2	A24407
20	14	25.0	10	2	C39745
21	14	25.0	10	2	D28027
22	14	25.0	10	2	S65715
23	14	25.0	10	2	A61218
24	14	25.0	10	2	B61218
25	14	25.0	10	2	A32195
26	14	25.0	10	2	E42364
27	13	23.2	5	2	B28495
28	13	23.2	9	2	S77984
29	13	23.2	9	2	S77984

30 13 23.2 9 2 PT0268 Ig heavy chain CRD  
31 13 23.2 10 2 S51912 hemagglutinin - in  
32 13 23.2 10 2 B45482 platelet activatin  
33 13 23.2 10 2 C39398 Fc mu (IgM) recept  
34 12 21.4 5 2 S11127 phosphoprotein, bo  
35 12 21.4 6 2 B34835 dnaA protein - Pse  
36 12 21.4 6 2 B26206 alpha-1,4-glucan-p  
37 12 21.4 6 2 A35039 hypothetical colla  
38 12 21.4 7 2 A61081 tryptophyllin, bas  
39 12 21.4 7 2 E61491 seed protein ws-5  
40 12 21.4 7 2 S42620 aggreacan - bovine  
41 12 21.4 7 2 PH1602 Ig H chain V-D-J r  
42 12 21.4 7 2 A39690 neural cell adhesi  
43 12 21.4 7 2 PH0932 T-cell receptor be  
44 12 21.4 8 2 S68802 nitrate reductase  
45 12 21.4 8 2 S10783 enamelini f - bovin

#### ALIGNMENTS

##### RESULT 1

I65546  
MHC H2-L antigen - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I65546

R:Kimura, A.; Israel, A.; Le Bail, O.; Kourilsky, P.

Cell 44, 261-272, 1986

A:Title: Detailed analysis of the mouse H-2Kb promoter: Enhancer-like sequences and their

A:Reference number: I52778; MUID:86106202; PMID:3510743

A:Accession: I65546

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6 <RES>

A:Cross-references: GB:M12483; NID:G199565; PIDN:AAA39663.1; PID:G554234

Query Match

Best Local Similarity 35.7%; Score 20; DB 2; Length 6;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VPC 10

Db 2 VPC 4

##### RESULT 2

A60521

glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)

N:Alternate names: glycogen phosphorylase b

C:Species: Liza ramada

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2003

C:Accession: A60521

R:Bonamusa, L.; Baanante, I.V.

Comp. Biochem. Physiol. B 95, 295-301, 1990

A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus

A:Reference number: A60521; MUID:90227907; PMID:2109669

A:Accession: A60521

A:Molecule type: protein

A:Residues: 1-5 <BON>

C:Superfamily: glucan phosphorylase

C:Keywords: glycosyltransferase; hexosyltransferase; phosphorylase b kinase) #status experim

F:3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase)

Query Match 33.9%; Score 19; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISVP 5

Db 2 ISVP 5

RESULT 3  
C39111  
IG heavy chain C region - Pacific hagfish (fragment)  
C/Species: Eptatretus stouti (Pacific hagfish)  
C/Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 16-Aug-1996  
C/Accession: C39111  
R/Varnier, J.; Neame, P.; Litman, G.W.  
Proc. Natl. Acad. Sci. U.S.A. 88, 1746-1750, 1991  
A/Title: A serum heterodimer from hagfish (Eptatretus stoutii) exhibits structural similarity  
A/Reference number: A39111; MUID:91156684; PMID:2000382  
A/Accession: C39111  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-10 <VAR>  
C/Keywords: heterotetramer; immunoglobulin

Query Match 33.9%; Score 19; DB 2; Length 10;  
Best Local Similarity 71.4%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISVPLSV 8  
|||  
DB 4 ISSPLV 10  
|||

RESULT 4  
PT0247  
IG heavy chain CRD3 region (clone 2-106A) - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C/Accession: PT0247  
R/Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A/Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A/Reference number: PT0222; MUID:91108337; PMID:1899102  
A/Accession: PT0247  
A/Molecule type: DNA  
A/Residues: 1-9 <YAM>  
A/Experimental source: B lymphocyte  
C/Keywords: heterotetramer; immunoglobulin

Query Match 32.1%; Score 18; DB 2; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 SVPLSV 9  
|||  
DB 2 SAPIDSP 8  
|||

RESULT 5  
A61622  
vitellogenin, 190k chain - gypsy moth (fragment)  
N/Contains: vitellin  
C/Species: Lymantria dispar (gypsy moth)  
C/Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 15-Oct-1999  
C/Accession: A61622  
R/Hiremath, S.; Eshita, S.  
Insect Biochem. Mol. Biol. 22, 605-611, 1992  
A/Title: Purification and characterization of vitellogenin from the gypsy moth, Lymantria  
A/Reference number: A61622  
A/Accession: A61622  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-10 <HIR>  
C/Keywords: egg yolk; hemolymph

Query Match 30.4%; Score 17; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLSVP 9  
|||

Db 2 PLTEP 6

# RESULT 6

PH0944  
T-cell receptor beta chain V-D-J region (clone 15) - rat (fragment)  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C/Accession: PH0944  
R/Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A/Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A/Reference number: PH0891; MUID:92078857; PMID:1836012  
A/Accession: PH0944  
A/Molecule type: mRNA  
A/Residues: 1-10 <GOL>

A/Experimental source: complete Freund's adjuvant-immunized lymph node  
A/Note: the authors translated the codon GAC for residue 9 as Glu  
C/Keywords: T-cell receptor

Query Match 30.4%; Score 17; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 4.8e+03;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CISVP 5  
|||  
DB 1 CASSP 5  
|||

# RESULT 7

I51049  
metallothionein-A - rainbow trout (fragment)  
C/Species: Oncorhynchus mykiss (rainbow trout)  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C/Accession: I51049  
R/Olsson, P.E.; Kling, P.; Erkel, L.J.; Kille, P.  
Eur. J. Biochem. 230, 344-349, 1995  
A/Title: Structural and functional analysis of the rainbow trout (Oncorhynchus mykiss)  
A/Reference number: I51049; MUID:95324545; PMID:7601121  
A/Accession: I51049  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-4 <OLS>  
A/Cross-references: EMBL:X80181; NID:g1019799; PIDN:CAA56466.1; PID:g4379328

Query Match 28.6%; Score 16; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PC 10  
|||  
DB 3 PC 4  
|||

# RESULT 8

G33098  
20SK exoantigen - malaria parasite (Plasmodium falciparum) (fragments)  
C/Species: Plasmodium falciparum  
C/Date: 24-Aug-1990 #sequence\_revision 24-Aug-1990 #text\_change 09-Jun-2000  
C/Accession: G33098  
R/Nichols, J.H.; Hager, L.P.  
submitted to the Protein Sequence Database, May 1990  
A/Reference number: A33098  
A/Accession: G33098  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-8 <NIC>

Query Match 28.6%; Score 16; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPLSV 8

Db 2 VFLXL 6  
|||:

## RESULT 9

S62880  
polygalacturonase (EC 3.2.1.15) IV - Aspergillus sp. (fragment)  
C:Species: Aspergillus sp.  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S62880  
R:Stratilova, E.; Dzurava, M.; Markovic, O.; Joernvall, H.  
FEBS Lett. 382, 164-166, 1996  
A:Title: An essential tyrosine residue of Aspergillus polygalacturonase.  
A:Reference number: S62880; MUID:96196586; PMID:8612742  
A:Accession: S62880  
A:Molecule type: protein  
A:Residues: 1-10 <STR>  
A:Keywords: glycosidase; hydrolase  
F:4/Active site: Tyr #status predicted

Query Match 28.6%; Score 16; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.2e+03; Mismatches 0; Indels 0; Gaps 0;  
Matches 2; Conservative 0;

QY 9 PC 10  
|||

Db 9 PC 10

## RESULT 10

PH0926  
T-cell receptor beta chain V-D-J region (isolate 12) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0926  
R:Gold, D.P.; Ofner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A:Reference number: PH0891; MUID:92078857; PMID:1836012  
A:Accession: PH0926  
A:Molecule type: mRNA  
A:Residues: 1-10 <GOL>  
A:Experimental source: concanavalin A-activated lymphoblast  
A:Note: the authors translated the codon AGA for residue 4 as Thr  
C:Keywords: T-cell receptor

Query Match 28.6%; Score 16; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 7.2e+03; Mismatches 2; Indels 0; Gaps 0;  
Matches 3; Conservative 0;

QY 1 CJSVP 5  
|||

Db 1 CASRP 5

## RESULT 11

A42689  
major postsynaptic density protein - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 26-May-1994  
C:Accession: A42689  
R:Wu, K.; Huang, Y.; Adler, J.; Black, I.B.  
Proc. Natl. Acad. Sci. U.S.A. 89, 3015-3019, 1992  
A:Title: On the identity of the major postsynaptic density protein.  
A:Reference number: A42689; MUID:92212958; PMID:1313576  
A:Accession: A42689  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <WUA>

Query Match 26.8%; Score 15; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05; Mismatches 3; Conservative 0; Indels 1; Gaps 0;

QY 6 LSVSP 9  
|||

Db 2 LXVP 5

## RESULT 12

B49712  
ATP-binding protein p46 - dog (fragment)  
C:Species: Canis lupus familiaris (dog)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: B49712  
R:Nigam, S.K.; Goldberg, A.L.; Ho, S.; Rohde, M.F.; Bush, K.T.; Sherman, M.Y.  
J. Biol. Chem. 269, 1744-1749, 1994  
A:Title: A set of endoplasmic reticulum proteins possessing properties of molecular chap  
A:Reference number: A49712; MUID:94124514; PMID:8294423  
A:Accession: B49712  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <NIG>  
C:Keywords: ATP; endoplasmic reticulum; molecular chaperone

Query Match 26.8%; Score 15; DB 2; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2.8e+05; Mismatches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 VPLSVP 9  
|||

Db 1 IPXXYP 6

## RESULT 13

S43630  
cytochrome-c oxidase (EC 1.9.3.1) chain VIc, hepatic - rainbow trout (fragment)  
C:Species: Oncorhynchus mykiss (rainbow trout)  
C:Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C:Accession: S43630  
R:Freund, R.; Kadenbach, B.  
Eur. J. Biochem. 221, 1111-1116, 1994  
A:Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytochro  
A:Reference number: S43624; MUID:94237150; PMID:8181469  
A:Accession: S43630  
A:Molecule type: protein  
A:Residues: 1-10 <PRE>  
A:Note: the source is designated as Salmo gairdneri  
C:Genetics:

A:Genome: nuclear  
C:Keywords: liver; membrane-associated complex; mitochondrion; oxidoreductase

Query Match 26.8%; Score 15; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.1e+04; Mismatches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 LSVSP 9  
|||

Db 2 LXVP 5

## RESULT 14

E86128  
hypothetical protein Z5903 [imported] - Escherichia coli (strain O157:H7, substrain ED19  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: E86128  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: E86128  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-10 <STO>

A;Cross-references: GB:AE005174; NID:g12519314; PIDN:AAG59489.1; GSPDB:GN00145; UWGP:Z59  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z5903

Query Match 26.8%; Score 15; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.1e+04;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 LSVP 10  
| | |  
Db 5 LQVVC 9

## RESULT 15

B22565

R-phycoerythrin alpha-2 chain - red alga (Gastroclonium coulteri) (fragment)

C;Species: Gastroclonium coulteri

C;Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993

C;Accession: B22565

R;Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A;Reference number: A22565; MUID:85182601; PMID:3886644

A;Accession: B22565

A;Molecule type: protein

A;Residues: 1-5 &lt;KLO&gt;

Query Match 25.0%; Score 14; DB 2; Length 5;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 LSVP 9  
| | |  
Db 1 LCVP 4

Search completed: September 5, 2004, 11:35:13  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:29:57 ; Search time 14 Seconds  
(without alignments)

37.193 Million cell updates/sec

Title: US-09-761-636A-13

Perfect score: 56

Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 371

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	46.4	9	1	MGMT_BOVIN
2	19	33.9	10	1	LPK2_LOCOMI
3	18	32.1	8	1	COM2_CONPU
4	15	26.8	10	1	COXH_ONCMY
5	15	26.8	10	1	NSL_MYCTU
6	15	26.8	10	1	TKL2_LOCOMI
7	15	26.8	10	1	TPF7_PACDA
8	14	25.0	7	1	TRP7_LEUMA
9	14	25.0	10	1	E101_LITRU
10	13	23.2	6	1	CONO_CONST
11	13	23.2	9	1	COXE_THUOB
12	13	23.2	9	1	COMQ_SHEEP
13	13	23.2	10	1	MNP1_LEPDE
14	12	21.4	7	1	ALL6_CYPDO
15	12	21.4	8	1	FUSS_FUSSO
16	12	21.4	8	1	UPA1_HUMAN
17	12	21.4	8	1	FARS_PENMO
18	12	21.4	9	1	FLA2_TREHY
19	12	21.4	9	1	OXYT_BUFRE
20	12	21.4	9	1	RT33_BOVIN
21	12	21.4	9	1	SAP_STOVA
22	12	21.4	9	1	ANGT_CHICK
23	12	21.4	10	1	COXQ_RABIT
24	12	21.4	10	1	UPA8_HUMAN
25	11	19.6	6	1	CIP1_MYTED
26	11	19.6	6	1	CIP2_MYTED
27	11	19.6	6	1	VP19_HSV1K
28	11	19.6	6	1	CARP_MYTED
29	11	19.6	7	1	ALL5_CRDPO
30	11	19.6	8	1	PPK3_PERAM
31	11	19.6	8	1	VGLG_HSV2B
32	11	19.6	8	1	DNF1_LOCOMI
33	11	19.6	9	1	

34 11 19.6 9 1 FARB\_ASCSU P43172 ascaris suu  
35 11 19.6 9 1 LMT3\_LOCOMI P41489 locusta mig  
36 11 19.6 9 1 OXYA\_SQUAC P42999 squalus aca  
37 11 19.6 9 1 OXYT\_RABIT P32878 oryctolagus  
38 11 19.6 9 1 UPA3\_HUMAN P30089 homo sapien  
39 11 19.6 9 1 UPA7\_HUMAN P30093 homo sapien  
40 11 19.6 10 1 AH3\_PRUSE P29261 prunus sero  
41 11 19.6 10 1 BPP8\_BOTIN P30426 bothrops in  
42 11 19.6 10 1 BPP\_VIPAS P31351 vipera aspi  
43 11 19.6 10 1 BRK\_ONCMY Q9PRZ1 oncorhynchu  
44 11 19.6 10 1 COXK\_ONCMY P80332 oncorhynchu  
45 11 19.6 10 1 GS09\_BACSU P80243 bacillus su

## ALIGNMENTS

RESULT 1  
MGMT\_BOVIN STANDARD; PRT; 9 AA.  
AC P29177;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Methylated-DNA-protein-cysteine methyltransferase (EC 2.1.1.63) (6-O-methylguanine-DNA methyltransferase) (Fragment).  
DE methylguanine-DNA methyltransferase (Fragment).  
GN MGMT.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae; Bovidae; Bovinae; Bos.  
OC NCBI\_TaxID=9913;  
RN [1]\_TaxID=9913;  
RP SEQUENCE.  
RC TISSUE=Thymus;  
RX MEDLINE=90174912; PubMed=2308822;  
RA Rydberg B., Hall J., Karran P.;  
RT "Active site amino acid sequence of the bovine O6-methylguanine-DNA methyltransferase.";  
RL Nucleic Acids Res. 18:17-21(1990).  
CC -!- FUNCTION: Repair of alkylated guanine in DNA by stoichiometrically transferring the alkyl group at the O-6 position to a cysteine residue in the enzyme. This is a suicide reaction: the enzyme is irreversibly inactivated.  
CC -!- CATALYTIC ACTIVITY: DNA (containing 6-O-methylguanine) + protein [protein]-L-cysteine = DNA (without 6-O-methylguanine) + protein S-methyl-L-cysteine.  
CC -!- SIMILARITY: WITH SEGMENTS OF E. COLI ADA AND OGT METHYLTRANSFERASE WHICH ENCOMPASS THE ALKYL-ACCEPTOR RESIDUES.  
DR InterPro: IPR001497; Methyltransf\_1.  
DR PROSITE: PS00374; MGMT; PARTIAL.  
KW DNA repair; Transferase; Methyltransferase.  
FT NON\_TER 1 1  
FT ACT\_SITE 9 9 ALKYL GROUP ACCEPTOR (BY SIMILARITY).  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 967 MW; 325171A720476047 CRC64;

Query Match 46.4%; Score 26; DB 1; Length 9;  
Best Local Similarity 42.9%; Pred. No. 1.4e+05;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPLSVPC 10  
Db :|:|  
3 IPILTPC 9

RESULT 2  
LPK2\_LOCOMI STANDARD; PRT; 10 AA.  
ID LPK2\_LOCOMI  
AC P41488;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Locustapyrokinin 2 (LOW-PK-2) (FXPRL-amide).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Brain;  
 RX MEDLINE=94094539; PubMed=7903606;  
 RA Schoofs L., Holman G.M., Nachman R., Proost P., van Damme J.,  
 RA de Loof A.;  
 RT "Isolation, identification and synthesis of locustapyrokinin II from  
 RT Locusta migratoria, another member of the FXPRL-amide peptide  
 RT family".  
 RL Comp. Biochem. Physiol. 106C:103-109(1993).  
 CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
 CC activity).  
 CC -1- SIMILARITY: Belongs to the pyrokinin family.  
 DR InterPro; IPR001484; PYROKININ.  
 DR PROSITE; PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin; Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1145 MW; CFAF4271A9D1B772 CRC64;  
 Query Match 33.9%; Score 19; DB 1; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 8.8e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 SVPLSVP 9  
 |||||  
 DB 2 SVTFTP 8

RESULT 3  
 COW2 CONPU STANDARD; PRT; 8 AA.  
 AC P58785;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Leu-contryphan-P.  
 OS Conus purpurascens (Purple cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=41690;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
 RC STRAIN=Clipperton Island; TISSUE=Venom;  
 RX MEDLINE=99388839; PubMed=10461743;  
 RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,  
 RA Olvera B.M.;  
 RT "A novel D-leucine-containing Conus peptide: diverse conformational  
 RT dynamics in the contryphan family.";  
 RL J. Pept. Res. 54:93-99(1999).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -1- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.  
 CC -1- SIMILARITY: Belongs to the contryphan family.  
 KW Toxin; Hydroxylation; D-amino acid.  
 FT DISULFID 2 8  
 FT MOD\_RES 4 4 D-LEUCINE.  
 FT MOD\_RES 4 4  
 SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;  
 Query Match 32.1%; Score 18; DB 1; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.4e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CISVP 5  
 |::|  
 DB 2 CVLLP 6

DE Locustapyrokinin 2 (LOW-PK-2) (FXPRL-amide).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Brain;  
 RX MEDLINE=94094539; PubMed=7903606;  
 RA Schoofs L., Holman G.M., Nachman R., Proost P., van Damme J.,  
 RA de Loof A.;  
 RT "Isolation, identification and synthesis of locustapyrokinin II from  
 RT Locusta migratoria, another member of the FXPRL-amide peptide  
 RT family".  
 RL Comp. Biochem. Physiol. 106C:103-109(1993).  
 CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
 CC activity).  
 CC -1- SIMILARITY: Belongs to the pyrokinin family.  
 DR InterPro; IPR001484; PYROKININ.  
 DR PROSITE; PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin; Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1145 MW; CFAF4271A9D1B772 CRC64;  
 Query Match 33.9%; Score 19; DB 1; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 8.8e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 SVPLSVP 9  
 |||||  
 DB 2 SVTFTP 8

RESULT 3  
 COW2 CONPU STANDARD; PRT; 8 AA.  
 AC P58785;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Leu-contryphan-P.  
 OS Conus purpurascens (Purple cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=41690;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
 RC STRAIN=Clipperton Island; TISSUE=Venom;  
 RX MEDLINE=99388839; PubMed=10461743;  
 RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,  
 RA Olvera B.M.;  
 RT "A novel D-leucine-containing Conus peptide: diverse conformational  
 RT dynamics in the contryphan family.";  
 RL J. Pept. Res. 54:93-99(1999).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -1- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.  
 CC -1- SIMILARITY: Belongs to the contryphan family.  
 KW Toxin; Hydroxylation; D-amino acid.  
 FT DISULFID 2 8  
 FT MOD\_RES 4 4 D-LEUCINE.  
 FT MOD\_RES 4 4  
 SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;  
 Query Match 32.1%; Score 18; DB 1; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.4e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CISVP 5  
 |::|  
 DB 2 CVLLP 6

RESULT 4  
 COXH ONCMY STANDARD; PRT; 10 AA.  
 ID -COXH ONCMY  
 AC P80331;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Cytochrome c oxidase polypeptide Vic (EC 1.9.3.1) (Fragment).  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Liver;  
 RX MEDLINE=94237150; PubMed=8181469;  
 RA Freund R., Kadenbach B.;  
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa  
 RT of cytochrome c oxidase isolated from rainbow trout.";  
 RL Eur. J. Biochem. 221:1111-1116(1994).  
 CC -1- FUNCTION: This protein is one of the nuclear-coded polypeptide  
 CC chains of cytochrome c oxidase, the terminal oxidase in  
 CC mitochondrial electron transport.  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H2O.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 DR PIR; S43630; S43630.  
 KW Oxidoreductase; Inner membrane; Mitochondrion.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 977 MW; E11B40769DC772DA CRC64;  
 Query Match 26.8%; Score 15; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 4.5e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 LSVP 9  
 |||||  
 DB 2 LXVP 5

RESULT 5  
 NS1 MYCTU STANDARD; PRT; 10 AA.  
 ID -NS1 MYCTU  
 AC P81135;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 30 kDa non-secretory protein 1 (Fragment).  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=H37Rv;  
 RA Prasad H.K., Annapurna P.S.;  
 RL Submitted (DEC-1997) to Swiss-Prot.  
 CC -1- CAUTION: We are unable to find this protein in the translation of  
 CC the genome of strain H37Rv.  
 FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1042 MW; 8767FE6AB2C73771 CRC64;  
 Query Match 26.8%; Score 15; DB 1; Length 10;  
 Best Local Similarity 37.5%; Pred. No. 4.5e+03;  
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 ISVPLSVP 9  
 ::|||  
 DB 1 MATPLVDP 8

## RESULT 6

TKL2\_LOCFI  
ID TKL2\_LOCFI STANDARD; PRT; 10 AA.  
AC P16224;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Locustatachykinin II (TX-II).  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=90184489; PubMed=2311766;  
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;  
RT "Locustatachykinin I and II, two novel insect neuropeptides with  
homology to peptides of the vertebrate tachykinin family.";  
RL FEBS Lett. 261:397-401(1990).  
CC -!- FUNCTION: Myoactive peptide. Stimulates the contraction of the  
ovoiduct and foregut.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
DR PIR; S08266; ECLQ2M.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
AMIDATION.  
SQ SEQUENCE 10 AA; 1066 MW; 5D52CD6B59C865B7 CRC64;

Query Match 26.8%; Score 15; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.5e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLS 7

DB 2 PLS 4

## RESULT 7

UPA2\_HUMAN  
ID UPA2\_HUMAN STANDARD; PRT; 10 AA.  
AC P30088;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Unknown protein from 2D-page of plasma (Spot 10) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=93092937; PubMed=1459097;  
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
RA Hochstrasser D.F.;  
RT "Plasma protein map: an update by microsequencing.";  
RL Electrophoresis 13:707-714(1992).  
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown  
protein is: 4.4, its MW is: 49 kDa.  
CC SWISS-2DPAGE; P30088; HUMAN.  
DR NON\_TER 1  
FT UNSURE 6 6  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1079 MW; 51AC54AAB77775B7 CRC64;

Query Match 26.8%; Score 15; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 4.5e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPLSVP 9  
DB 1 VXLSP 6

## RESULT 8

TPFY\_PACDA  
ID TPFY\_PACDA STANDARD; PRT; 7 AA.  
AC P83455;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tryptophyllin-1 (pdt-1)  
OS Pachymedusa dactylosa (Giant mexican leaf frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;  
OC Phyllomedusinae; Pachymedusa.  
OX NCBI\_TaxID=75988;  
RN [1]  
RP SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF  
PRO-7.  
RC TISSUE=Skin secretion;  
RA Chen T.B., Orr D.F., Shaw C.;  
RT "Pachymedusa dactylosa tryptophyllin-1 (pdt-1): structural  
characterization, pharmacological activity and cloning of precursor  
cDNA.";  
RL Submitted (SEP-2002) to Swiss-Prot.  
CC -!- FUNCTION: Myoactive. Has selective relaxing activity on vascular  
smooth muscle.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Skin.  
CC -!- MASS SPECTROMETRY: MW=809.2; METHOD=WALDI.  
DR GO; GO:000576; C:extracellular; NAS.  
DR GO; GO:0045986; P:negative regulation of smooth muscle contra. . .; NAS.  
KW Amphibian defense peptide; Amidation; Hydroxylation.  
FT MOD\_RES 3 3  
HYDROXYLATION.  
FT MOD\_RES 7 7  
AMIDATION.  
SQ SEQUENCE 7 AA; 794 MW; 7772D37DC776350 CRC64;

Query Match 25.0%; Score 14; DB 1; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+05;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLSVP 9

DB 3 PAWVP 7

## RESULT 9

TRP7\_LEUMA  
ID TRP7\_LEUMA STANDARD; PRT; 10 AA.  
AC P81739;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tachykinin-related peptide 7 (LemTRP 7).  
OS Leucophaea madeira (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
OC Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Brain;  
RX MEDLINE=97269266; PubMed=9114447;  
RA Muren J.E., Naessel D.R.;  
RT "Seven tachykinin-related peptides isolated from the brain of the  
madeira cockroach; evidence for tissue-specific expression of  
isoforms.";  
RL Peptides 18:7-15(1997).  
CC -!- FUNCTION: Myoactive peptide. Increases the amplitude and frequency  
of spontaneous contractions and tonus of hindgut muscle.

```

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Brain.
CC -1- MASS SPECTROMETRY: MW=1069.7; METHOD=WALDI.
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
FW Tachykinin; Neuropeptide; Amidation.
KT MOD RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1068 MW; C4541679C9C865BD CRC64;

Query Match 25.0%; Score 14; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.7e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPLS 7
DB 1 VPAS 4

RESULT 10
EIO1 LITRU ID EIO1 LITRU STANDARD; PRT; 6 AA.
AC P82056;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 23.2%; Score 13; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPL 6
DB 2 VPI 4

RESULT 11
CONO CONST ID CONO CONST STANDARD; PRT; 9 AA.
AC P05487;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arg-conopressin S.
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;
RN [1]
RP SEQUENCE.
RX MEDLINE=88058932; PubMed=3680228;
RA Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeikus R.D.,
RA Gray W.R., Olivera B.M.;
RT "Invertebrate vasopressin/oxytocin homologs. Characterization of
RT peptides from Conus geographus and Conus striatus venoms.";
```

```

RL J. Biol. Chem. 262:15821-15824(1987).
RN [2]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -1- FUNCTION: Targets vasopressin-oxytocin related receptors.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- SIMILARITY: Belongs to the vasopressin/oxytocin family.
DR PIR; B28495; B28495.
DR InterPro; IPR000981; Neurhyp_horm.
DR Pfam; PF00220; hormone4; 1.
DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1031 MW; 17EB176EB4540050 CRC64;

Query Match 23.2%; Score 13; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CI 2
DB 1 CI 2

RESULT 12
COXE THUOB ID COXE THUOB STANDARD; PRT; 9 AA.
AC P80975;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome c oxidase polypeptide Via (BC 1.9.3.1) (Fragment).
OS Thunus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -1- CATALYTIC ACTIVITY: 4 ferrocyanide c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase Via family.
DR PIR; S77984; S77984.
DR InterPro; IPR001349; COX6A.
DR PROSITE; PS01329; COX6A; PARTIAL.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 1 1
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1136 MW; 62E072C9CB0776DB CRC64;

Query Match 23.2%; Score 13; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLSVP 9
```



```

Db          4 PEFVP 8
RESULT 13
COXQ SHEEP
ID COXQ SHEEP STANDARD; PRT; 10 AA.
AC P80337;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIII-Liver/heart (EC 1.9.3.1)
DE (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RA Freund R., Kadenbach B.;
RL Submitted (MAR-1994) to Swiss-Prot.
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase VIII family.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1027 MW; C4E95CA33DC7633D CRC64;
Query Match 23.2%; Score 13; DB 1; Length 10;
Best Local Similarity 40.0%; Pred. No. 1e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLSVP 9
Db 5 PAKTP 9
RESULT 14
MNPI LEPDE
ID MNPI LEPDE STANDARD; PRT; 7 AA.
AC P42984;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Myotropic neuropeptide 1 (Led-MNP-I).
OS Leptinotarsa decemlineata (Colorado potato beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phycophaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;
OC Chrysomelini; Leptinotarsa.
OX NCBI_TaxID=7539;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Spittaels K., Vankeerberghen A., Schoofs L., Torrekens S.,
RA Grauwels L., van Leuven F., de Loof A.;
RT "Identification, characterization, and immunological localization of
RT a novel myotropic neuropeptide in the Colorado potato beetle,
RT Leptinotarsa decemlineata."
RL Peptides 16:365-374(1995).
CC -!- FUNCTION: Myotropic peptide. Stimulates the contractions of the
CC oviduct.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 705 MW; 6DD73768745B5DB0 CRC64;
Query Match 21.4%; Score 12; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLS 7
Db 5 PLA 7
RESULT 15
ALU6 CYDPO
ID ALU6 CYDPO STANDARD; PRT; 8 AA.
AC P82157;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastomina 6.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: Belongs to the allatostatin family.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 936 MW; 0B2879C45B573767 CRC64;
Query Match 21.4%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPL 6
Db 1 LPL 3
Search completed: September 5, 2004, 11:33:40
Job time : 14 secs

```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:30:23 ; Search time 57 seconds  
(without alignments)  
55.354 Million cell updates/sec

Title: US-09-761-636A-13  
Perfect score: 56  
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1443

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	35.7	9	11 Q35953	Q35953 mus musculus
2	20	35.7	10	15 Q85598	Q85598 moloney mur
3	20	35.7	10	15 Q85563	Q85563 moloney mur
4	20	35.7	10	15 Q85619	Q85619 moloney mur
5	19	33.9	8	2 Q56140	Q56140 streptococc
6	19	33.9	10	2 Q7WUG2	Q7WUG2 escherichia
7	17	30.4	8	4 Q9Y4X6	Q9Y4X6 homo sapien
8	17	30.4	9	10 Q7X6A3	Q7X6A3 zea mays su
9	17	30.4	10	8 Q96041	Q96041 oenothera b
10	17	30.4	10	10 Q7Y0I8	Q7Y0I8 zea mays su
11	17	30.4	10	12 Q39957	Q39957 hepatitis g
12	16	28.6	8	4 Q15898	Q15898 homo sapien
13	16	28.6	8	6 Q02831	Q02831 oryctolagus
14	16	28.6	8	6 Q9TRY3	Q9TRY3 sus sp. ins
15	16	28.6	9	12 Q8QVD3	Q8QVD3 ovine respi
16	16	28.6	9	13 Q8AYL5	Q8AYL5 carassius a

17	16	28.6	9	13 Q8AUM7	Q8AUM7 carassius a
18	16	28.6	10	6 Q8SPN8	Q8SPN8 macaca mula
19	16	28.6	10	8 Q8SAZ9	Q8SAZ9 pyrrhobryum
20	16	28.6	10	10 Q94IS6	Q94IS6 pinus taeda
21	15	26.8	7	13 Q42564	Q42564 fugu rubrip
22	15	26.8	10	2 Q9L5W6	Q9L5W6 liberibacte
23	15	26.8	10	16 Q8X4E5	Q8X4E5 escherichia
24	14	25.0	7	15 Q07624	Q07624 rous sarcom
25	14	25.0	8	2 Q9X3K1	Q9X3K1 prochloroco
26	14	25.0	9	2 P72345	P72345 pseudomonas
27	14	25.0	9	2 Q9K4M6	Q9K4M6 staphylococ
28	14	25.0	9	4 Q9UCS8	Q9UCS8 homo sapien
29	14	25.0	9	5 Q9TUV0	Q9TUV0 anthopieura
30	14	25.0	9	10 Q9FXL0	Q9FXL0 liliun long
31	14	25.0	10	4 Q9H3R9	Q9H3R9 homo sapien
32	14	25.0	10	4 Q9H121	Q9H121 homo sapien
33	14	25.0	10	8 Q8MAZ9	Q8MAZ9 dicranostyl
34	14	25.0	10	8 Q8MBB7	Q8MBB7 merremia ae
35	14	25.0	10	8 Q8SL54	Q8SL54 aeonium haw
36	14	25.0	10	10 P83091	P83091 spinacia ol
37	14	25.0	10	12 Q90347	Q90347 hepatitis g
38	14	25.0	10	12 Q8JV78	Q8JV78 polyomaviru
39	14	25.0	10	13 Q9PRU1	Q9PRU1 cynops pyrr
40	13	23.2	7	12 Q67113	Q67113 influenzavi
41	13	23.2	8	2 Q8KXP4	Q8KXP4 microcystis
42	13	23.2	8	2 Q49534	Q49534 mycoplasma
43	13	23.2	8	2 Q32560	Q32560 escherichia
44	13	23.2	8	4 Q8IUB8	Q8IUB8 homo sapien
45	13	23.2	8	8 Q8WFR5	Q8WFR5 diadema pau

#### ALIGNMENTS

RESULT 1

O35953 PRELIMINARY; PRT; 9 AA.  
AC O35953;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).  
GN SCNBA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RII1;  
RX MEDLINE=97442476; PubMed=9295353;  
RA Plummer N.W., McBurney M.W., Meisler M.H.;  
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated two-domain protein in fetal brain and non-neuronal cells.";  
RL J. Biol. Chem. 272:24008-24015(1997).  
DR EMBL; U97672; AAB80914.1; -;  
DR MGD; MGI:103169; Scn8a.  
DR GO; GO:0007628; P:adult walking behavior; IMP.  
KW Ionic channel.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 898 MW; 22D92865B735B737 CRC64;

Query Match 35.7%; Score 20; DB 11; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPLSV 8

Db 1 VPLSL 5

RESULT 2

Q85598

```

ID Q85598 PRELIMINARY; PRT; 10 AA.
AC Q85598;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Moloney murine sarcoma virus (Strain HT1) env/mos 5' junction
DE (Fragment).
DE Moloney murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11801;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164305; PubMed=6300424;
RA Donoghue D.J., Hunter T.;
RT "Recombination junctions of variants of Moloney murine sarcom virus:
RT Generation and divergence of a mammalian transforming gene.";
RL J. Virol. 45:607-617(1983).
DR EMBL; K03106; AAA46492.1; -.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1081 MW; 7BECFCBEA771B5A4 CRC64;

Query Match 35.7%; Score 20; DB 15; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SVPC 10
Db |||
4 STPC 7

RESULT 3
Q85563 PRELIMINARY; PRT; 10 AA.
ID Q85563;
AC Q85563;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Env-mos fusion protein (Fragment).
OS Moloney murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11801;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82196891; PubMed=6281735;
RA Donoghue D.J., Hunter T.;
RT "A generalized method of subcloning DNA fragments by restriction site
RT reconstruction: Application to sequencing the amino-terminal region of
RT the transforming gene of Gazdar murine sarcoma virus.";
RL Nucleic Acids Res. 10:2549-2564(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164305; PubMed=6300424;
RA Donoghue D.J., Hunter T.;
RT "Recombination junctions of variants of Moloney murine sarcom virus:
RT Generation and divergence of a mammalian transforming gene.";
RL J. Virol. 45:607-617(1983).
DR EMBL; K03105; AAA46491.1; -.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1062 MW; F9ECFCBEA771B5B1 CRC64;

Query Match 35.7%; Score 20; DB 15; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SVPC 10
Db |||
4 STPC 7

RESULT 4
Q85619 PRELIMINARY; PRT; 10 AA.
ID Q85619;
AC Q85619;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Moloney murine sarcoma virus (Strain ml) env/mos 5' junction
DE (Fragment).
OS Moloney murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11801;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164305; PubMed=6300424;
RA Donoghue D.J., Hunter T.;
RT "Recombination junctions of variants of Moloney murine sarcom virus:
RT Generation and divergence of a mammalian transforming gene.";
RL J. Virol. 45:607-617(1983).
DR EMBL; K03108; AAA46494.1; -.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1081 MW; 7BECFCBEA771B5A4 CRC64;

Query Match 35.7%; Score 20; DB 15; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SVPC 10
Db |||
4 STPC 7

RESULT 5
Q56140 PRELIMINARY; PRT; 8 AA.
ID Q56140;
AC Q56140;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE STP6 protein (Fragment).
GN STP6.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ST11;
RX MEDLINE=95047254; PubMed=7958782;
RA Constable A., Mollet B.;
RT "Isolation and characterisation of promoter regions from Streptococcus
RT thermophilus.";
RL FEMS Microbiol. Lett. 122:85-90(1994).
DR EMBL; X78210; CAA55045.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 846 MW; ED086772D5B045B6 CRC64;

Query Match 33.9%; Score 19; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISVP 5
Db |||
3 ISVP 6

RESULT 6
Q7WUG2 PRELIMINARY; PRT; 10 AA.
ID Q7WUG2;
AC Q7WUG2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TnpIS1 (Fragment).
GN INSA.
OS Escherichia coli.
OG Plasmid p541.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Miriagou V., Tzouveleakis L.S., Villa L., Lebessi E., Vatsopoulos A.C.,
RA Carattoli A., Tzelepi E.;
RT "Antibiotic Resistance Region of an IncN plasmid Carrying an Integron-
RT Located blaVIM-1-Metallo- $\beta$ -Lactamase Gene and a Novel CMY-Type
RT Cephalosporinase Gene.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY339625; AAQ16673.1; -.
KW Plasmid.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 991 MW; 882D57A5B045A2D5 CRC64;

Query Match 33.9%; Score 19; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. No. 5.8e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SVPLSVP 9
Db |||:|
3 SVSISCP 9

RESULT 7
Q9Y4X6 PRELIMINARY; PRT; 8 AA.
ID Q9Y4X6;
AC Q9Y4X6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nuclear LIM interactor (Fragment).
GN NLI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=2010806; PubMed=10640831;
RA Drechsler M., Schumacher V., Friedrich S., Wildhardt G., Giesler S.,
RA Schroth A., Bodem J., Royer-Pokora B.;
RT "Genomic structure, alternative transcripts and chromosome location of
RT the human LIM domain binding protein gene LDB1.";
RL Cytogenet. Cell Genet. 87:119-124 (1993).
DR EMBL; A0243097; CAB45408.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 767 MW; EE6BBDDEB862D5B6 CRC64;

Query Match 30.4%; Score 17; DB 4; Length 8;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 LSVPC 10
Db |||
1 MSVGC 5

RESULT 8
Q7X6A3 PRELIMINARY; PRT; 9 AA.
ID Q7X6A3;
AC Q7X6A3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Isoamylase (Fragment).
GN SU1.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. 38-11, and cv. A632;
RA Wilson L.M., Whitt S.R., Ibanez A.M., Rocheford T.R., Goodman M.M.,
RA Buckler E.S. IV.;
RT "Dissection of maize starch production by candidate gene
RT association.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY290305; AAP45331.1; -.
DR EMBL; AY290311; AAP45337.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 976 MW; DF9BCEA76736C6DD CRC64;

Query Match 30.4%; Score 17; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VPC 10
Db |||
5 LPC 7

RESULT 9
Q96041 PRELIMINARY; PRT; 10 AA.
ID Q96041;
AC Q96041;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH-ubiquinone oxidoreductase subunit 3 (Fragment).
GN NAD3.
OS Oenothera lutea (Bertero's evening primrose).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=3950;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94019250; PubMed=8413195;
RA Schuster W.;
RT "Ribosomal protein gene rpl5 is cotranscribed with the nad3 gene in
RT Oenothera mitochondria.";
RL Mol. Gen. Genet. 240:445-449 (1993).
DR EMBL; X69553; CAA49285.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion; Ubiquinone.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1097 MW; 723067B0476DD9CB CRC64;

Query Match 30.4%; Score 17; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CIS 3
Db |||
8 CIS 10

RESULT 10
Q7Y018 PRELIMINARY; PRT; 10 AA.
ID Q7Y018;
AC Q7Y018;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Isoamylase (Fragment).
GN SU1.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=cv. Kl44;
RC  Wilson L.M., Whitt S.R., Ibanez A.M., Rocheford T.R., Goodman M.M.,
RA  Buckler E.S. IV.;
RT  "Dissection of maize starch production by candidate gene
RT  association.";
RL  Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY290360; AAF45386.1; -.
FT  NON_TER 10
SQ  SEQUENCE 10 AA; 1063 MW; D0FF9BCEA76736C6 CRC64;

Query Match 30.4%; Score 17; DB 10; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.4e+04;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VPC 10
Db 5 LPC 7

RESULT 11
ID 039957 PRELIMINARY; PRT; 10 AA.
AC 039957;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE E1 protein (Fragment).
OS Hepatitis GB virus C.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC GBV-C/HGV group.
OX NCBI_TaxID=39839;
RN [1]

SEQUENCE FROM N.A.
RP STRAIN=Edinburgh haemophilic;
RX MEDLINE=97368412; PubMed=9225026;
RA Smith D.B., Cuccane N., Davidson F., Jarvis L.M., Mokili J.L.,
RA Hamid S., Ludlam C.A., Simmonds P.;
RT "Discrimination of hepatitis G virus/GBV-C geographical variants by
RT analysis of the 5' non-coding region.";
RL J. Gen. Virol. 78:1533-1542(1997).
DR EMBL; AF003175; AAC57986.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1067 MW; CC88FE27273772 CRC64;

Query Match 30.4%; Score 17; DB 12; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.4e+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ISVPL 6
Db 1 MAVPL 5

RESULT 12
ID Q15898 PRELIMINARY; PRT; 8 AA.
AC Q15898;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP6A11B) (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.R., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;

```

```

RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32078; AAA73888.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 938 MW; 34A415B0477B45BB CRC64;

Query Match 28.6%; Score 16; DB 4; Length 8;
Best Local Similarity 60.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLS 7
Db 2 SYFIS 6

RESULT 13
ID 002831 PRELIMINARY; PRT; 8 AA.
AC 002831;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pro alpha 1 type III collagen protein (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]

SEQUENCE FROM N.A.
RP MEDLINE=96377339; PubMed=8783186;
RA Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA Vuorio E.;
RT "Evidence for insufficient chondrocytic differentiation during repair
RT of full-thickness defects of articular cartilage.";
RL Matrix Biol. 15:39-47(1996).
DR EMBL; S83371; AAD14433.1; -.
KW Collagen.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

Query Match 28.6%; Score 16; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PC 10
Db 3 PC 4

RESULT 14
ID Q9TRY3 PRELIMINARY; PRT; 8 AA.
AC Q9TRY3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Insulin-like growth factor-binding protein-6, IGFBP-6 (Fragment).
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]

SEQUENCE.
RP MEDLINE=92049376; PubMed=1719383;
RA Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
RT protein-6.";
RL Mol. Endocrinol. 5:938-948(1991).
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 850 MW; 9FB2CEA37EA7687D CRC64;

```

Query Match 28.6%; Score 16; DB 6; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1e+06; 0;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PC 10  
||  
Db 3 PC 4

## RESULT 15

Q8QVD3 PRELIMINARY; PRT; 9 AA.  
AC Q8QVD3;  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Matrix protein 2 (Fragment).  
GN M2.  
OS Ovine respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_taxID=28869;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21580659; PubMed=11724134;  
RA Eleraky N.Z., Kania S.A., Potgieter L.N.;  
RT "The ovine respiratory syncytial virus F gene sequence and its  
RT diagnostic application."  
RL J. Vet. Diagn. Invest. 13:455-461(2001).  
DR EMBL; AF334398; AAL91343.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1154 MW; 8B6A3EA764541415 CRC64;

Query Match 28.6%; Score 16; DB 12; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1e+06;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PC 10  
||  
Db 6 PC 7

Search completed: September 5, 2004, 11:34:44  
Job time : 58 secs

***This Page Blank (uspio)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:27:12 ; Search time 63 Seconds

(without alignments)  
44.849 Million cell updates/sec

Title: US-09-761-636A-13

Perfect score: 56

Sequence: 1 CISVPLSVC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 325896

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	100.0	10	4	Aau04532 VEGF base
2	41.5	74.1	9	4	Aau04533 VEGF base
3	36	64.3	9	2	Aar96138 Protease
4	36	64.3	9	2	Aaw82212 D-NoirFES-
5	36	64.3	9	2	Aaw46562 Peptide b
6	33	58.9	9	2	Aar96137 Protease
7	33	58.9	9	2	Aaw82094 D-NoirFES-
8	33	58.9	9	2	Aaw46561 Peptide b
9	33	58.9	9	4	Aag73245 Protease
10	33	58.9	9	5	Abu60357 D-NoirFES-
11	32	57.1	10	4	Aau00643 Human mem
12	32	57.1	10	7	Adb88786 Membrane
13	31.5	56.2	9	5	Aau94301 Human nov
14	31.5	56.2	10	5	Aau94201 Human nov
15	31.5	56.2	10	5	Aau94811 Human nov
16	31.5	56.2	10	5	Aau94577 Human nov
17	31	55.4	10	5	Aam50003 Human D40
18	30	53.6	9	5	Aau95231 Human nov
19	30	53.6	9	5	Aau94696 Human nov
20	30	53.6	9	5	Aau95023 Human nov
21	29	51.8	10	4	Aag73418 Human gen
22	29	51.8	10	4	Aag85638 Saccharom
23	29	51.8	10	5	Abg64264 Human alb
24	28	50.0	9	6	Abri9062 Human can
25	28	50.0	10	6	Abri9524 Human can

## ALIGNMENTS

RESULT 1  
ID AAU04532 standard; peptide; 10 AA.  
XX  
XX AAU04532;  
AC  
DT 26-SEP-2001 (first entry)  
XX  
DE VEGF based monocyclic peptide 10.  
XX  
KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
KW diabetic retinopathy; chronic inflammation; cyclic.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..10  
FT  
XX  
PN WO200152875-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 18-JAN-2001; 2001WO-US001533.  
XX  
PR 18-JAN-2000; 2000US-0176293P.  
PR 16-MAY-2000; 2000US-0204590P.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Achen MG, Hughes RA, Stacker S, Cendron A;  
XX WPI; 2001-442248/47.  
XX  
PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
PT residues.  
XX  
PS Claim 49; Page 32; 102pp; English.  
XX  
CC The sequence represents a monomeric monocyclic peptide of the invention,  
CC whose 3-dimensional structure is modelled on the expose loop of human  
CC VEGF (vascular endothelial growth factor). The invention relates to a  
CC method of producing a monomeric monocyclic peptide by a measuring beta-  
CC beta carbon separation distances on opposite antiparallel strands of a

Abri8877 Human can  
Aar74541 Protease  
Adc71003 HLA motif  
Adc70887 HLA motif  
Adc70636 HLA motif  
Adc70959 HLA motif  
Adc70732 HLA motif  
Adc70653 HLA motif  
Adc71049 HLA motif  
Aap82102 Non-label  
Aag93495 Human bon  
Abj04424 Stem cell  
Aam48672 CXCR4 pep  
Aaw12561 SH2 bindi  
Aau96540 Human cyt  
Aau04529 VEGF base  
Aar77368 SH3 bindi  
Aag98618 Human cel  
Aae05003 Human rel  
Aab60620 Human MUM

CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 4; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0098;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 C1SVPLSVPC 10  
 |||||  
 Db 1 C1SVPLSVPC 10

RESULT 2  
 ID AAU04533  
 AAU04533 standard; peptide; 9 AA.

AC AAU04533;  
 DT 26-SEP-2001 (first entry)  
 DE VEGF based monocyclic peptide 11.  
 XX Human; VEGF; vascular endothelial growth factor; angiogenesis;  
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
 KW diabetic retinopathy; chronic inflammation; cyclic.  
 XX  
 OS Synthetic.

Key Location/Qualifiers  
 FT Disulfide-bond 1..9 /note= "This bond cyclises the peptide"  
 FT

XX WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US001533.

XX 18-JAN-2000; 2000US-0176293P.

XX 16-MAY-2000; 2000US-0204590P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
 PT or lymphangiogenesis, is produced by cyclising a peptide loop fragment  
 FT from an exposed loop of a growth factor protein by oxidizing the cysteine

PT residues.

XX Claim 49; Page 32; 102pp; English.

CC The sequence represents a monomeric monocyclic peptide of the invention,  
 CC whose 3-dimensional structure is modelled on the exposed loop of human  
 CC VEGFD (vascular endothelial growth factor). The invention relates to a  
 CC method of producing a monomeric monocyclic peptide by a measuring beta-  
 CC beta carbon separation distances on opposite antiparallel strands of a  
 CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 9 AA;

Query Match 74.1%; Score 41.5; DB 4; Length 9;  
 Best Local Similarity 90.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 C1SVPLSVPC 10  
 |||||  
 Db 1 C1SVPL-VPC 9

RESULT 3

AA96138

ID AAR96138 standard; peptide; 9 AA.

XX AAR96138;

XX 25-MAR-2003 (revised)

DT 18-DEC-1996 (first entry)

XX Protease substrate peptide with fluorophore at each terminus.

XX Fluorogenic substrate; fluorophore; protease activity; assay;

XX visible fluorescence; in situ detection; frozen tissue section;

XX histology; arthritis; emphysema; thrombosis; cancer metastasis.

XX Synthetic.

XX Key Location/Qualifiers

XX Region 1..3

XX /label= C1

XX /note= "the two conformation determining regions C1 and

XX C2 which flank the protease recognition site peptide P

XX are provided to position the two fluorophores within 100

XX angstroms of each other"

XX Modified-site 1

XX /note= "labelled by donor fluorophore (F1) 5'-

XX carboxytetramethylrhodamine"

XX 4..7

XX /label= P

XX /note= "peptide comprising a protease recognition site"

FT Region 8...9  
 FT /label= C2  
 FT /note= "the two conformation determining regions C1 and  
 FT C2 which flank the protease recognition site peptide P  
 FT are provided to position the two fluorophores within 100  
 FT angstroms of each other"  
 FT 9  
 FT Modified-site  
 FT /note= "labelled by acceptor fluorophore F2 rhodamine X  
 FT acetamide"  
 FT  
 PN W09613607-A1.  
 XX  
 XX  
 PD 09-MAY-1996.  
 XX  
 PD 27-OCT-1995; 95WO-US013936.  
 PF  
 XX 28-OCT-1994; 94US-00331383.  
 PR  
 XX (ONCO-) ONCOIMMUNIN INC.  
 PA  
 PI Komoriya A, Packard BS;  
 XX  
 XX WPI; 1996-239512/24.  
 DR  
 XX New fluorogenic peptide(s) with fluorophore at each terminus - for  
 PT detecting protease(s) in biological samples, emit intense visible  
 PT fluorescence when cleaved.  
 XX  
 PS Claim 16; Page 31; 89pp; English.  
 XX  
 CC The present sequence is a specific example of a fluorogenic substrate for  
 CC detecting activity of a protease. The substrate agrees with the generic  
 CC formula (S1)n-C1(F1)-P-C2(F2)-(S2)k in which a peptide P of 2-8 amino  
 CC acids comprising a recognition site for the protease is flanked by  
 CC conformation determining regions (C1 and C2) of 1-3 amino acids. C1 and  
 CC C2 are labelled by fluorophore groups (F1 and F2, respectively)  
 CC positioned within 100 angstroms of each other. Additional peptide spacers  
 CC of 1-50 amino acids (S1 and S2) may also be present (i.e. n and k = 0 or  
 CC 1). Fluorogenic substrates corresponding to the generic formula are used  
 CC to detect or localise proteases in biological specimens, esp. in frozen  
 CC tissue sections or to monitor protease activity in stored reagents.  
 CC Changes in protease activity are associated with e.g. arthritis,  
 CC emphysema, thrombosis and cancer metastasis. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 PS Sequence 9 AA;  
 XX  
 CC Query Match 64.3%; Score 36; DB 2; Length 9;  
 CC Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 CC Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 DB ::||:|  
 2 AIPMSGPC 9  
 RESULT 4  
 AAW82212  
 ID AAW82212 standard; peptide; 9 AA.  
 XX  
 AC AAW82212;  
 XX  
 XX 18-FEB-1999 (first entry)  
 DT  
 XX D-NorFES-A protease inhibitor peptide #2.  
 DE  
 XX Protease activity; fluorophore; detection; fluorogenic; cellular uptake;  
 KW conformation change.  
 XX  
 XX Synthetic.  
 OS  
 XX W09837226-A1.  
 PN  
 XX

PD 27-AUG-1998.  
 XX  
 PF 20-FEB-1998; 98WO-US003000.  
 XX  
 PR 20-FEB-1997; 97US-00802981.  
 XX  
 PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Komoriya A, Packard BS;  
 XX  
 XX WPI; 1998-467579/40.  
 DR  
 XX New fluorogenic compositions - containing 2 fluorophores separated by a  
 PT peptide comprising a protease binding site, used for detecting protease  
 PT activity in samples.  
 XX  
 PS Example 1; Page 52; 90pp; English.  
 XX  
 CC AAW82023-W82240 are peptides used in the construction of a fluorogenic  
 CC composition which is used for the detection of protease activity in  
 CC biological samples. The products can be used for the detection of  
 CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,  
 CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,  
 CC steroids or polymers. In addition, attachment of a hydrophobic group to a  
 CC molecule can be used to enhance uptake by cells. The composition is  
 CC composed of P = peptide comprising a protease binding site for the  
 CC protease, F1, F2 peptides = fluorophores where F1 is attached to the  
 CC amino terminal amino acid and F2 is attached to the carboxyl terminal  
 CC amino acid and S1, S2 peptides = when present, are peptide spacers where  
 CC S1, when present, is attached to the amino terminal acid, and S2, when  
 CC present, is attached to the carboxyl terminal amino acid  
 XX  
 PS Sequence 9 AA;  
 XX  
 CC Query Match 64.3%; Score 36; DB 2; Length 9;  
 CC Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 CC Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 DB ::||:|  
 2 AIPMSGPC 9  
 RESULT 5  
 AAW45562  
 ID AAW45562 standard; peptide; 9 AA.  
 XX  
 AC AAW45562;  
 XX  
 DT 20-MAY-1998 (first entry)  
 XX  
 XX Peptide backbone of a protease indicator.  
 DE  
 XX Protease binding site; protease; protease indicator; fluorescent signal;  
 KW detection; protease activity.  
 XX  
 OS Synthetic.  
 XX  
 PN US5714342-A.  
 PN  
 PD 03-FEB-1998.  
 XX  
 XX 27-OCT-1995; 95US-00549008.  
 PF  
 XX 28-OCT-1994; 94US-00331383.  
 XX  
 PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Packard BS, Komoriya A;  
 XX  
 XX WPI; 1998-158345/14.  
 DR  
 XX Fluorogenic substrates for protease determination - having two closely  
 PT

FT spaced fluorophores flanking protease binding site.  
 XX  
 PS Example 1; Col 23; 39pp; English.  
 XX  
 CC The present peptide contains a protease binding site. It is used to  
 CC produce novel reagents whose fluorescence increases in the presence of  
 CC particular proteases. These fluorogenic protease indicators (substrates)  
 CC provide a high intensity fluorescent signal at a visible wavelength when  
 CC they are digested by a protease. The fluorogenic indicators have the  
 CC general formula: F1-C1-P-C2-F2 | (S1)n (S2)k where: P is a peptide  
 CC containing a protease binding site e.g. AAW46520-53, AAW46560. F1 and F2  
 CC are fluorophores. S1 and S2 are peptide spacers e.g. AAW46554-58. n, k =  
 CC 0 or 1. C1 and C2 are conformation-determining regions that introduce a  
 CC bend into the composition which positions the fluorophores adjacent to  
 CC each other with a separation of less than 100 Angstrom. When n is 1, S1  
 CC is joined to the terminal alpha -amino group of C1 by a peptide bond, and  
 CC when k is 1, S2 is joined to the terminal carboxy group of C2 by a  
 CC peptide bond. The protease indicators are used for detecting protease  
 CC activity in a biological sample. The sample is contacted with the  
 CC indicator and any change in fluorescence is detected, an increase in  
 CC fluorescence indicating protease activity  
 XX  
 SQ Sequence 9 AA;  
 Query Match 64.3%; Score 36; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 Db :||:|  
 2 AIPMSIPC 9  
 RESULT 6  
 AAR96137  
 ID AAR96137 standard; peptide; 9 AA.  
 XX  
 AC AAR96137;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 18-DEC-1996 (first entry)  
 XX  
 DE Protease substrate peptide with fluorophore at each terminus.  
 XX  
 KW Fluorogenic substrate; fluorophore; protease activity; assay;  
 KW visible fluorescence; in situ detection; frozen tissue section;  
 KW histology; arthritis; emphysema; thrombosis; cancer metastasis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..3  
 FT /label= C1  
 FT /note= "the two conformation determining regions C1 and  
 FT C2 which flank the protease recognition site peptide P  
 FT are provided to position the two fluorophores within 100  
 FT angstroms of each other"  
 FT Modified-site 1  
 FT /note= "labelled by donor fluorophore (F1) 5'-  
 FT carboxytetramethylrhodamine"  
 FT Peptide 4..7  
 FT /label= P  
 FT /note= "peptide comprising a protease recognition site"  
 FT Modified-site 5  
 FT /label= Nle  
 FT Region 8..9  
 FT /note= "the two conformation determining regions C1 and  
 FT C2 which flank the protease recognition site peptide P  
 FT are provided to position the two fluorophores within 100  
 FT angstroms of each other"  
 FT Modified-site 9  
 FT /note= "labelled by acceptor fluorophore F2 rhodamine X

FT acetamide"  
 XX WO9613607-A1.  
 PN  
 XX  
 PD 09-MAY-1996.  
 XX  
 PF 27-OCT-1995; 95WO-US013936.  
 XX  
 PP 28-OCT-1994; 94US-00331383.  
 PR  
 XX (ONCO-) ONCOIMMUNIN INC.  
 PA Komoriya A, Packard BS;  
 PI WPI; 1996-239512/24.  
 DR  
 XX New fluorogenic peptide(s) with fluorophore at each terminus - for  
 PT detecting protease(s) in biological samples, emit intense visible  
 PT fluorescence when cleaved.  
 XX  
 PS Claim 15; Page 31; 88pp; English.  
 PS  
 XX The present sequence is a specific example of a fluorogenic substrate for  
 CC detecting activity of a protease. The substrate agrees with the generic  
 CC formula (S1)n-C1(F1)-P-C2(F2)-(S2)k in which a peptide P of 2-8 amino  
 CC acids comprising a recognition site for the protease is flanked by  
 CC conformation determining regions (C1 and C2) of 1-3 amino acids. C1 and  
 CC C2 are labelled by fluorophore groups (F1 and F2, respectively)  
 CC positioned within 100 angstroms of each other. Additional peptide spacers  
 CC of 1-50 amino acids (S1 and S2) may also be present (i.e. n and k = 0 or  
 CC 1). Fluorogenic substrates corresponding to the generic formula are used  
 CC to detect or localise proteases in biological specimens, esp. in frozen  
 CC tissue sections or to monitor protease activity in stored reagents.  
 CC Changes in protease activity are associated with e.g. arthritis, emphysema,  
 CC thrombosis and cancer metastasis. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 SQ Sequence 9 AA;  
 Query Match 58.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 Db :||:|  
 2 AIPXSIPC 9  
 RESULT 7  
 AAW82094  
 ID AAW82094 standard; peptide; 9 AA.  
 XX  
 AC AAW82094;  
 XX  
 DT 18-FEB-1999 (first entry)  
 DT  
 XX D-NorFES-A protease inhibitor peptide.  
 DE  
 XX Protease activity; fluorophore; detection; fluorogenic; cellular uptake;  
 KW conformation change.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "Asp residue is modified by the presence of F1  
 FT where F1 is the donor fluorophore 5'-  
 FT carboxytetramethylrhodamine (C2211)"  
 FT Modified-site 5  
 FT /label= Nle  
 FT /note= "Norleucine"  
 FT Modified-site 9  
 FT /note= "Cys residue is modified by the presence of F2

FT FT where F2 is the acceptor fluorophore rhodamine X  
 FT acetamide (R492)"  
 PN  
 XX WO9837226-A1.  
 XX  
 XX 27-AUG-1998.  
 PD  
 XX  
 PF 20-FEB-1998; 98WO-US003000.  
 XX  
 XX 20-FEB-1997; 97US-00802981.  
 PR  
 XX (ONCO-) ONCOIMMUNIN INC.  
 XX  
 XX Komoriya A, Packard BS;  
 PI  
 XX WPI; 1998-467579/40.  
 DR  
 XX  
 XX New fluorogenic compositions - containing 2 fluorophores separated by a  
 PT peptide comprising a protease binding site, used for detecting protease  
 PT activity in samples.  
 XX  
 XX Disclosure; Page 10; 90pp; English.  
 PS  
 XX AAW2023-W82240 are peptides used in the construction of a fluorogenic  
 CC composition which is used for the detection of protease activity in  
 CC biological samples. The products can be used for the detection of  
 CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,  
 CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,  
 CC steroids or polymers. In addition, attachment of a hydrophobic group to a  
 CC molecule can be used to enhance uptake by cells. The composition is  
 CC composed of P = peptide comprising a protease binding site for the  
 CC protease, F1, F2 peptides = fluorophores where F1 is attached to the  
 CC amino terminal amino acid and F2 is attached to the carboxyl terminal  
 CC amino acid and S1, S2 peptides = when present, are peptide spacers where  
 CC S1, when present, is attached to the amino terminal acid, and S2, when  
 CC present, is attached to the carboxyl terminal amino acid  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 58.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 Db ::||:|  
 2 AIPXSIPC 9  
 RESULT 8  
 AAW46561  
 ID AAW46561 standard; peptide; 9 AA.  
 XX  
 AC AAW46561;  
 XX  
 XX 20-MAY-1998 (first entry)  
 DT  
 XX Peptide backbone of a protease indicator.  
 DE  
 XX  
 XX Protease binding site; protease; protease indicator; fluorescent signal;  
 KW detection; protease activity.  
 XX  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 PH Modified-site 5 /label= Nle  
 FT /note= "Norleucine"  
 FT  
 XX US5714342-A.  
 PN  
 XX  
 XX 03-FEB-1998.  
 PD  
 XX  
 XX 27-OCT-1995; 95US-00549008.  
 PF

XX 28-OCT-1994; 94US-00331383.  
 PR  
 XX (ONCO-) ONCOIMMUNIN INC.  
 PA  
 XX Packard BS, Komoriya A;  
 PI  
 XX WPI; 1998-158345/14.  
 DR  
 XX  
 XX Fluorogenic substrates for protease determination - having two closely  
 PT spaced fluorophores flanking protease binding site.  
 PT  
 XX Example 1; Col 23; 39pp; English.  
 PS  
 XX The present peptide contains a protease binding site. It is used to  
 CC produce novel reagents whose fluorescence increases in the presence of  
 CC particular proteases. These fluorogenic protease indicators (substrates)  
 CC provide a high intensity fluorescent signal at a visible wavelength when  
 CC they are digested by a protease. The fluorogenic indicators have the  
 CC general formula: F1--C1--P--C2--F2 | (S1)n (S2)k where: P is a peptide  
 CC containing a protease binding site e.g. AAW46520-53, AAW46560, F1 and F2  
 CC are fluorophores. S1 and S2 are peptide spacers e.g. AAW46554-58. n, k =  
 CC 0 or 1. C1 and C2 are conformation-determining regions that introduce a  
 CC bend into the composition which positions the fluorophores adjacent to  
 CC each other with a separation of less than 100 Angstrom. When n is 1, S1  
 CC is joined to the terminal alpha -amino group of C1 by a peptide bond, and  
 CC when k is 1, S2 is joined to the terminal carboxy group of C2 by a  
 CC peptide bond. The protease indicators are used for detecting protease  
 CC activity in a biological sample. The sample is contacted with the  
 CC indicator and any change in fluorescence is detected, an increase in  
 CC fluorescence indicating protease activity  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 58.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 Db ::||:|  
 2 AIPXSIPC 9  
 RESULT 9  
 AAG73245  
 ID AAG73245 standard; peptide; 9 AA.  
 XX  
 AC AAG73245;  
 XX  
 XX 14-AUG-2001 (first entry)  
 DT  
 XX Protease indicator peptide #1.  
 DE  
 XX  
 XX Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis; thrombosis;  
 KW haemophilia.  
 XX  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /label= OTHER  
 FT /note= "modified by fluorophore"  
 FT  
 FT Modified-site 9 /label= OTHER  
 FT /note= "modified by fluorophore"  
 FT  
 XX WO200118238-A1.  
 PN  
 XX  
 XX 15-MAR-2001.  
 PD  
 XX  
 XX 11-SEP-2000; 2000WO-US024882.  
 PF  
 XX

PR 10-SEP-1999; 99US-00394019.  
 XX (ONCO-) ONCOIMMUNIN INC.  
 PA Komoriya A, Packard BS;  
 XX WPI; 2001-389573/41.  
 DR  
 XX  
 XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples.  
 XX  
 XX Example 2; Page 53; 86pp; English.  
 PS  
 XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 9 AA;  
 Query Match 58.9%; Score 33; DB 4; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 :|||:  
 Db 2 AIPXSIPC 9  
 RESULT 10  
 ABU60357  
 ID ABU60357 standard; peptide; 9 AA.  
 XX  
 AC ABU60357;  
 XX  
 XX 29-APR-2003 (first entry)  
 DE D-NorFES-A protease inhibitor peptide.  
 XX  
 XX Protease; indicator; chromophore; H-dimer; fluorescence; absorbance;  
 KW nuclease; screening; fluorophore; substrate cleavage.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "F1, where F1 is a donor fluorophore 5'-  
 FT carbocytetramethylrhodamine (C2211)"  
 FT Modified-site 5  
 FT /label= Nle  
 FT /note= "norleucine"  
 FT Modified-site 9  
 FT /note= "F2, where F2 is an acceptor fluorophore rhodamine  
 FT X acetamide (R492)"  
 XX  
 XX WO200261038-A2.  
 PN  
 XX  
 XX 08-AUG-2002.  
 PD  
 XX  
 XX 21-DEC-2001; 2001WO-US049781.  
 PF  
 XX  
 XX 22-DEC-2000; 2000US-00747287.  
 PR  
 XX (ONCO-) ONCOIMMUNIN INC.  
 PA  
 XX Packard BS, Komoriya A;  
 PI  
 XX

DR WPI; 2002-698548/75.  
 XX  
 PT Indicator composition comprising polypeptide or nucleic acid backbone  
 PT joining two same chromophores resulting in quenching of fluorescence  
 PT of/change in absorbance of chromophores, useful for detecting protease  
 PT activity.  
 XX  
 XX Example 2; Page 15; 97pp; English.  
 PS  
 XX This invention describes a novel indicator composition (referred as homo-  
 CC doubly labeled compositions) comprising a polypeptide backbone or a  
 CC nucleic acid backbone joining two chromophores of the same species  
 CC whereby the chromophores form an H-dimer resulting in quenching of the  
 CC fluorescence of or a change in the absorbance of the chromophore, a  
 CC decrease in fluorescence or a change in absorbance indicates that the  
 CC first molecule and the second molecule are interacting. The indicator is  
 CC useful for detecting the activity of a protease, where an increase in  
 CC fluorescence or a change in absorbance indicates that the protease  
 CC cleaves the polypeptide backbone. The indicator is attached to a solid  
 CC support inside a mammalian, yeast or insect cell. The composition bears a  
 CC hydrophobic group such as Fmoc, 9-fluoreneacetyl group, 1-  
 CC fluorene-carboxylic group, 9-fluorene-carboxylic group, and 9-fluorenone-1-  
 CC carboxylic group, benzyloxycarbonyl, Xanthyl (Xan), Trityl (Trt), 4-  
 CC methyltrityl (Mtt), 4-methoxytrityl (Mmt), 4-methoxy-2,3, 6-trimethyl-  
 CC benzenesulphonyl (Mtr), Mesitylene-2-sulphonyl (Mts), 4,4'-  
 CC dimethoxybenzhydryl (Mbh), etc. The method described in the invention is  
 CC useful for detecting protease or nuclease activity (or the presence of  
 CC nucleic acid) in histological section, cells in culture, (e.g., seeded or  
 CC cultured adherent cells), a biological sample such as tissue, biopsy,  
 CC lymph, embryo, or whole animal, or cell suspension derived from a  
 CC biological sample such as tissue, blood, urine, saliva, lymph, or biopsy.  
 CC The indicator composition is also useful for screening a test agent for  
 CC the ability to modulate a protease (or a nuclease, lipase, etc.). The  
 CC indicator reagents allow rapid determination of protease activity in a  
 CC matter of minutes in a single-step procedure. The fluorescent indicators  
 CC both absorb and emit in the visible range (400-800 nm). These signals are  
 CC therefore not readily quenched by, nor is activation of the fluorophores,  
 CC that is, absorption of light, interfered with by background molecules;  
 CC therefore they are easily detected in biological samples. The fluorogenic  
 CC protease indicators utilize high efficiency fluorophores and are able to  
 CC achieve a high degree of quenching while providing a strong signal when  
 CC the quench is released by cleavage of the peptide substrate. The high  
 CC signal allows detection of very low levels of protease activity. Thus the  
 CC fluorogenic protease indicators are particularly well suited for in situ  
 CC detection of protease activity. ABU60357-ABU60477 represent peptides use  
 CC to illustrate the method described in the disclosure of the invention  
 XX  
 SQ Sequence 9 AA;  
 Query Match 58.9%; Score 33; DB 5; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SVPLSVPC 10  
 :|||:  
 Db 2 AIPXSIPC 9  
 RESULT 11  
 AAU00643  
 ID AAU00643 standard; peptide; 10 AA.  
 XX  
 XX AAU00643;  
 AC  
 XX  
 XX 07-SEP-2001 (first entry)  
 DT  
 XX Human membrane translocating peptide (MTLP) #12.  
 DE  
 XX Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
 KW epithelial cell layer; gastrointestinal tract; circulatory system.  
 XX  
 XX Homo sapiens.  
 XX







02-JUL-2002 (first entry)  
Human novel protein CaTrF2E11 HLA binding peptide #394.  
Human; human leukocyte antigen; HLA; immunogen; 83P2H3; CaTrF2E11;  
calcium transport protein; cancer; prostate cancer; cytostatic;  
chromosome 7q34; chromosome 12q24.1; T cell; B cell.  
Homo sapiens.  
WO200214361-A2.  
21-FEB-2002.  
17-AUG-2001; 2001WO-US025782.  
17-AUG-2000; 2000US-0226329P.  
(AGEN-) AGENSYS INC.  
Raitano AB, Challita-Eid PM, Faris M, Saffran DC, Afar DEH;  
Levin E, Hubert RS, Ge W, Jakobovits A;  
WPI; 2002-269179/31.  
Monitoring 83P2H3 gene products for monitoring the presence of cancer in  
a subject, comprises determining the status of 83P2H3 gene products in a  
tissue sample from the subject and comparing it to a normal sample.  
Example 11; Page 191; 270pp; English.  
The invention relates to monitoring 83P2H3 (a calcium transport protein  
whose gene is located on chromosome 7q34) gene products in a biological  
sample from a patient who has or is suspected of having cancer  
(especially prostate cancer), comprises: (a) determining the status of  
83P2H3 gene products expressed by cells in a tissue sample from an  
individual and (b) comparing the status to the status of 83P2H3 gene  
products in a normal sample. Also included are modulators of 83P2H3  
function or status, generating antibodies/immune response against 83P2H3  
for related protein CaTrF2E11 whose gene is located on chromosome  
12q24.1) using identified HLA (human leukocyte antigen) binding peptides  
derived from the protein, delivering a cytotoxic agent to a cell  
expressing 83P2H3 by conjugating the agent to an anti-83P2H3 antibody, a  
recombinant protein comprising an antigen-binding region of the antibody,  
a non-human transgenic animal that produces the recombinant protein, a  
hybridoma that produces the recombinant protein, a single-chain  
monoclonal antibody that comprises the variable domains of the heavy and  
light chains of the anti-83P2H3 antibody, a vector comprising a  
polynucleotide that encodes the monoclonal antibody and inducing an  
immune response to a 83P2H3 protein, by providing a 83P2H3-related  
protein that comprises a T cell or B cell epitope, and contacting the  
epitope with an immune system T cell or B cell, respectively. The method  
is useful for monitoring 83P2H3 gene products in a biological sample for  
monitoring the presence of cancer in an individual. The modulator is  
useful for inhibiting the growth of cancer cells that express 83P2H3, for  
treating cancer and the vector is useful for treating a patient with a  
cancer that expresses 83P2H3. The immunological methods are useful for  
generating an immune response against 83P2H3, and for detecting the  
presence of 83P2H3-related protein or polynucleotide in a biological  
sample from a patient who has or who is suspected of having cancer. The  
antibody is useful in prostate cancer diagnosis, prognosis, imaging  
methodologies and treatment, to detect and quantify 83P2H3 and mutant  
83P2H3-related proteins, for purifying a 83P2H3-related protein, for  
isolating 83P2H3 homologues/related molecules, and for generating anti-  
idiotypic antibodies that mimic the 83P2H3 protein. The present sequence  
is an HLA binding peptide motif from 83P2H3 or its related protein  
CaTrF2E11

Sequence 10 AA;

Query Match 56.2%; Score 31.5; DB 5; Length 10;  
Best Local Similarity 60.0%; Pred. No. 99;

Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
Qy 1 CTSVPLSVPC 10  
|::||| |  
Db 1 CLT-PLSFEC 9

Search completed: September 5, 2004, 11:33:20  
Job time : 64 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:34:49 ; Search time 66 Seconds  
(without alignments)  
47.737 Million cell updates/sec

Title: US-09-761-636A-13  
Perfect score: 56  
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 159066

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	9	US-09-761-636A-13
2	41.5	74.1	9	9	US-09-761-636A-14
3	33	58.9	9	12	US-09-747-287-1
4	33	58.9	9	12	US-09-874-350A-1
5	33	58.9	9	12	US-09-874-350A-184
6	32	57.1	10	14	US-10-126-845-13
7	32	57.1	10	14	US-10-126-845-71
8	32	57.1	10	15	US-10-116-275-101
9	32	57.1	10	16	US-10-764-235-13
10	31.5	56.2	9	10	US-09-932-165-284
11	31.5	56.2	10	10	US-09-932-165-184
12	31.5	56.2	10	10	US-09-932-165-560
13	31.5	56.2	10	10	US-09-932-165-794
14	31	55.4	9	15	US-10-154-884B-11221
15	31	55.4	10	12	US-10-363-791-194

16	53.6	9	10	US-09-932-165-679	Sequence 679, App
17	53.6	9	10	US-09-932-165-1006	Sequence 1006, App
18	53.6	9	10	US-09-932-165-1214	Sequence 1214, App
19	51.8	10	11	US-09-833-245-1011	Sequence 1011, App
20	48.2	8	9	US-09-756-283A-66	Sequence 66, Appl
21	48.2	9	15	US-10-154-884B-11228	Sequence 11228, A
22	48.2	9	15	US-10-154-884B-11250	Sequence 11250, A
23	48.2	9	15	US-10-154-884B-11254	Sequence 11254, A
24	48.2	9	15	US-10-154-884B-11258	Sequence 11258, A
25	48.2	9	15	US-10-154-884B-11259	Sequence 11259, A
26	48.2	9	15	US-10-154-884B-11265	Sequence 11265, A
27	48.2	9	15	US-10-154-884B-11275	Sequence 11275, A
28	48.2	9	15	US-10-154-884B-11280	Sequence 11280, A
29	48.2	9	16	US-10-415-014-421	Sequence 421, App
30	48.2	9	16	US-10-415-014-537	Sequence 537, App
31	48.2	10	16	US-10-415-014-187	Sequence 187, App
32	48.2	10	16	US-10-415-014-266	Sequence 266, App
33	48.2	10	16	US-10-415-014-493	Sequence 493, App
34	48.2	10	16	US-10-415-014-583	Sequence 583, App
35	48.2	9	9	US-09-852-424-32	Sequence 32, Appl
36	46.4	9	9	US-09-852-424-43	Sequence 43, Appl
37	46.4	9	12	US-10-363-208-21	Sequence 21, Appl
38	46.4	9	15	US-10-154-884B-11276	Sequence 11276, A
39	46.4	9	15	US-10-154-884B-11287	Sequence 11287, A
40	46.4	9	9	US-09-761-636A-10	Sequence 10, Appl
41	44.6	10	9	US-09-922-261-21	Sequence 21, Appl
42	42.9	5	9	US-09-866-135-5	Sequence 5, Appl
43	42.9	5	10	US-09-886-135-5	Sequence 5, Appl
44	42.9	9	12	US-09-935-430-12	Sequence 12, Appl
45	42.9	9	12	US-09-935-430-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-761-636A-13  
; Sequence 13, Application US/09761636A  
; Patent No. US20020065218A1  
; GENERAL INFORMATION:  
; APPLICANT: ACHEN, Marc  
; APPLICANT: STACKER, Steven  
; APPLICANT: HUGHES, Richard  
; APPLICANT: CENDRON, Angela  
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR  
; FILE REFERENCE: 1064/48505 Achen et al  
; CURRENT APPLICATION NUMBER: US/09761,636A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/176,293  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/204,590  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-761-636A-13

Query Match 100.0% ; Score 56; DB 9; Length 10;  
Best Local Similarity 100.0% ; Pred. No. 0.023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10  
|||  
Db 1 CISVPLSVPC 10

RESULT 2  
US-09-761-636A-14  
; Sequence 14, Application US/09761636A  
; Patent No. US20020065218A1

```

; GENERAL INFORMATION:
; APPLICANT: ACHEN, Marc
; APPLICANT: STACKER, Steven
; APPLICANT: HUGHES, Richard
; APPLICANT: CENDRON, Angela
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR
; FILE REFERENCE: 1064/48505 Achen et al
; CURRENT APPLICATION NUMBER: US/09/761,636A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/176,293
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: US 60/204,590
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-761-636A-14

```

```

Query Match          74.1%; Score 41.5; DB 9; Length 9;
Best Local Similarity 90.0%; Pred. No. 1.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY      1 C1SVPLSVPC 10
        ||||| |||
Db      1 C1SVPL-VPC 9

```

## RESULT 3

```

US-09-747-287-1
; Sequence 1, Application US/09747287
; Publication No. US20030207264A1
; GENERAL INFORMATION:
; APPLICANT: KOMORIYA, AKIRA
; TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME
; FILE REFERENCE: 300-948600US
; CURRENT APPLICATION NUMBER: US/09/747,287
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/349,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US08/802,981
; PRIOR FILING DATE: 1997-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide.
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Xaa is norleucine
US-09-747-287-1

```

```

Query Match          58.9%; Score 33; DB 12; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.2e+06;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      3 SVPLSVPC 10
        :||| |||
Db      2 AIPXSIPC 9

```

## RESULT 4

```

US-09-874-350A-1

```

```

; Sequence 1, Application US/09874350A
; Publication No. US20040096926A1
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903840US
; CURRENT APPLICATION NUMBER: US/09/874,350A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: US 09/394,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 221
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: protease indicator
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Xaa is norleucine (Nle)
US-09-874-350A-1

```

```

Query Match          58.9%; Score 33; DB 12; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.2e+06;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      3 SVPLSVPC 10
        :||| |||
Db      2 AIPXSIPC 9

```

## RESULT 5

```

US-09-874-350A-184
; Sequence 184, Application US/09874350A
; Publication No. US20040096926A1
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903840US
; CURRENT APPLICATION NUMBER: US/09/874,350A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: US 09/394,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 221
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 184
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide substrate
; NAME/KEY: MOD_RES
; LOCATION: (5)..(5)
; OTHER INFORMATION: X is norleucine

```

US-09-874-350A-184

Query Match 58.9%; Score 33; DB 12; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10  
Db 2 AIPXSIFC 9

## RESULT 6

US-10-126-845-13  
; Sequence 13, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126.845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic  
US-10-126-845-13

Query Match 57.1%; Score 32; DB 14; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10  
Db 1 CLPVLLAAPC 10

## RESULT 7

US-10-126-845-71  
; Sequence 71, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126.845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 71  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D form peptide  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: D form amino acid  
US-10-126-845-71

Query Match 57.1%; Score 32; DB 14; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10  
Db 1 CLPVLLAAPC 10

## RESULT 8

US-10-116-275-101  
; Sequence 101, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116.275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 101  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Listed in Table titled "Peptides that Target to and/or Enhance Uf  
; OTHER INFORMATION: take Across the GIT"  
US-10-116-275-101

Query Match 57.1%; Score 32; DB 15; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10  
Db 1 CLPVLLAAPC 10

## RESULT 9

US-10-764-235-13  
; Sequence 13, Application US/10764235  
; Publication No. US20040138132A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: P26,479-B USA  
; CURRENT APPLICATION NUMBER: US/10/764.235  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic  
US-10-764-235-13

Query Match 57.1%; Score 32; DB 16; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10

OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-184

Query Match 56.2%; Score 31.5; DB 10; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 C1SVPLSVPC 10  
|::|||  
Db 1 CLT-PLSFPC 9

RESULT 12  
US-09-932-165-560  
; Sequence 560, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: AEFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOVITS, AYA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CatrF2E11 USEFUL IN TREATMENT AND  
; TITLE OF INVENTION: DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 560  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-560

Query Match 56.2%; Score 31.5; DB 10; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 C1SVPLSVPC 10  
|::|||  
Db 1 CLT-PLSFPC 9

RESULT 13  
US-09-932-165-794  
; Sequence 794, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: AEFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOVITS, AYA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CatrF2E11 USEFUL IN TREATMENT AND  
; TITLE OF INVENTION: DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17

Db 1 CLPVLAAPC 10

RESULT 10  
US-09-932-165-284  
; Sequence 284, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: AEFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOVITS, AYA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CatrF2E11 USEFUL IN TREATMENT AND  
; TITLE OF INVENTION: DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 284  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-284

Query Match 56.2%; Score 31.5; DB 10; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.2e+06;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 C1SVPLSVPC 10  
|::|||  
Db 1 CLT-PLSFPC 9

RESULT 11  
US-09-932-165-184  
; Sequence 184, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: AEFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOVITS, AYA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CatrF2E11 USEFUL IN TREATMENT AND  
; TITLE OF INVENTION: DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 184  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:

; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 794  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-794

Query Match 56.2%; Score 31.5; DB 10; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 CISVPLSVPC 10  
|::|  
Db 1 CLT-PLSFPC 9

RESULT 14  
US-10-154-884B-11221  
; Sequence 11221, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11221  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-154-884B-11221

Query Match 55.4%; Score 31; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.2e+06;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CISVPLS 7  
|:|  
Db 2 CLSVFVS 8

RESULT 15  
US-10-363-791-194  
; Sequence 194, Application US/10363791  
; Publication No. US20040029197A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKIMOTO, Masato  
; APPLICANT: KUZUMAKI, No. US20040029197A1oru  
; APPLICANT: SATO, No. US20040029197A1iyuki  
; APPLICANT: SAHARA, Hiroeki  
; TITLE OF INVENTION: A novel human cancer/testis-associated gene thereof  
; FILE REFERENCE: 4439-4006  
; CURRENT APPLICATION NUMBER: US/10/363,791  
; CURRENT FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: JP 2000-274218  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 227  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 194  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-363-791-194

Query Match 55.4%; Score 31; DB 12; Length 10;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLSVPC 10  
|||  
Db 2 PLSAPC 7

Search completed: September 5, 2004, 11:40:12  
Job time : 67 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:32:18 ; Search time 15 Seconds  
(without alignments)  
34.417 Million cell updates/sec

Title: US-09-761-636A-13  
Perfect score: 56  
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 103740

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2.6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/prodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/prodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/prodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	64.3	9	1 US-08-331-383-33	Sequence 33, Appl
2	36	64.3	9	1 US-08-549-008-43	Sequence 43, Appl
3	36	64.3	9	3 US-08-802-981-143	Sequence 143, Appl
4	33	58.9	9	1 US-08-331-383-31	Sequence 31, Appl
5	33	58.9	9	1 US-08-549-008-42	Sequence 42, Appl
6	33	58.9	9	3 US-08-802-981-142	Sequence 142, Appl
7	25	44.6	10	1 US-08-212-190A-5	Sequence 5, Appl
8	25	44.6	10	2 US-08-900-321-5	Sequence 5, Appl
9	25	44.6	10	3 US-09-461-697-21	Sequence 21, Appl
10	25	44.6	10	5 PCT-US95-03610-5	Sequence 5, Appl
11	24	42.9	6	1 US-08-483-434A-21	Sequence 21, Appl
12	24	42.9	6	3 US-08-476-134A-30	Sequence 30, Appl
13	24	42.9	6	6 5190920-26	Patent No. 5190920
14	24	42.9	6	6 5506208-28	Patent No. 5506208
15	24	42.9	9	1 US-08-331-383-35	Sequence 35, Appl
16	24	42.9	9	1 US-08-549-008-40	Sequence 40, Appl
17	24	42.9	9	3 US-08-802-981-1	Sequence 1, Appl
18	24	42.9	10	1 US-08-033-857A-5	Sequence 5, Appl
19	24	42.9	10	3 US-08-374-983A-5	Sequence 17, Appl
20	24	42.9	10	3 US-08-377-781A-17	Sequence 17, Appl
21	23	41.1	6	1 US-07-994-133-3	Sequence 3, Appl
22	23	41.1	6	1 US-08-221-078A-8	Sequence 8, Appl
23	23	41.1	6	1 US-08-221-171A-8	Sequence 8, Appl
24	23	41.1	8	1 US-08-189-331-147	Sequence 147, Appl
25	23	41.1	8	1 US-08-189-331-148	Sequence 148, Appl
26	23	41.1	8	1 US-08-189-331-149	Sequence 149, Appl
27	23	41.1	8	1 US-08-189-331-150	Sequence 150, Appl

28	23	41.1	8	2 US-08-471-068-147	Sequence 147, App
29	23	41.1	8	2 US-08-471-068-148	Sequence 148, App
30	23	41.1	8	2 US-08-471-068-149	Sequence 149, App
31	23	41.1	8	2 US-08-471-068-150	Sequence 150, App
32	23	41.1	8	4 US-09-311-784A-371	Sequence 371, App
33	23	41.1	8	4 US-09-187-330-6	Sequence 6, Appli
34	23	41.1	9	4 US-09-997-579-19	Sequence 19, Appl
35	23	41.1	9	4 US-09-187-330-28	Sequence 28, Appl
36	23	41.1	10	4 US-09-187-330-33	Sequence 33, Appl
37	23	41.1	10	4 US-09-535-852-1686	Sequence 1686, Ap
38	22.5	40.2	9	3 US-08-891-271-4	Sequence 4, Appli
39	22.5	40.2	9	3 US-08-660-092-126	Sequence 126, App
40	22.5	40.2	9	4 US-09-160-513-126	Sequence 126, App
41	22	39.3	7	2 US-08-934-222-13	Sequence 13, Appl
42	22	39.3	7	2 US-08-933-402-13	Sequence 13, Appl
43	22	39.3	7	2 US-09-207-621-13	Sequence 13, Appl
44	22	39.3	7	2 US-08-532-818-13	Sequence 13, Appl
45	22	39.3	7	3 US-09-231-797-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-331-383-33  
; Sequence 33, Application US/08331383  
; Patent No. 5605809  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; TITLE OF INVENTION: Proteases for the Detection of  
; TITLE OF INVENTION: Proteases in Biological Samples and Methods and Use  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,383  
; FILING DATE: 28-OCT-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 16865-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-331-383-33

Query Match 64.3%; Score 36; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3e+05;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10  
Db 2 AIPMSIPC 9

```
RESULT 2
US-08-549-008-43
; Sequence 43, Application US/08549008
; Patent No. 5714342
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of
; TITLE OF INVENTION: Protease in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,008
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,383
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 016865-000110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-549-008-43

Query Match 64.3%; Score 36; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 2 AIPMSIPC 9

RESULT 3
US-08-802-981-143
; Sequence 143, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

Query Match 64.3%; Score 36; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 2 AIPMSIPC 9

RESULT 3
US-08-802-981-143
; Sequence 143, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

Query Match 64.3%; Score 36; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 2 AIPMSIPC 9

RESULT 4
US-08-331-383-31
; Sequence 31, Application US/08331383
; Patent No. 5605809
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of
; TITLE OF INVENTION: Proteases in Biological Samples and Methods and Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,383
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 16865-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-000300US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 143:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-802-981-143

Query Match 64.3%; Score 36; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 2 AIPMSIPC 9

RESULT 4
US-08-331-383-31
; Sequence 31, Application US/08331383
; Patent No. 5605809
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of
; TITLE OF INVENTION: Proteases in Biological Samples and Methods and Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,383
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 16865-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

;  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: one-of(5)  
; OTHER INFORMATION: /note= "Xaa is norleucine."  
US-08-331-383-31

Query Match 58.9%; Score 33; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3e+05; Length 9;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLSVPC 10  
:|:|:  
Db 2 AIPXSIPC 9

## RESULT 5

US-08-549-008-42  
; Sequence 42, Application US/08549008  
; Patent No. 5714342  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of  
; TITLE OF INVENTION: Protease in Biological Samples and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/549,008  
; FILING DATE: 27-OCT-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 08/331,383  
; APPLICATION NUMBER: 28-OCT-1994  
; FILING DATE: 28-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 016865-000110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 5  
; OTHER INFORMATION: /product= "Nle"

## US-08-549-008-42

Query Match 58.9%; Score 33; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLSVPC 10  
:|:|:  
Db 2 AIPXSIPC 9

## RESULT 6

US-08-802-981-142  
; Sequence 142, Application US/08802981  
; Patent No. 6037137  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of Enzyme  
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/802,981  
; FILING DATE: 20-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 016865-000300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 142:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 5  
; OTHER INFORMATION: /product= "Nle"

## US-08-802-981-142

Query Match 58.9%; Score 33; DB 3; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLSVPC 10  
:|:|:  
Db 2 AIPXSIPC 9

## RESULT 7

US-08-212-190A-5  
; Sequence 5, Application US/08212190A  
; Patent No. 5652223  
; GENERAL INFORMATION:  
; APPLICANT: KOHN, Elise C.  
; APPLICANT: LIOTTA, Lance A.  
; APPLICANT: KIM, Young Sook  
; TITLE OF INVENTION: DNA ENCODING CAI RESISTANCE PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California

QY 3 SVPLSVPC 10  
:|:|:  
Db 2 AIPXSIPC 9

```
;
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; APPLICATION DATA:
; FILING DATE: 14-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 15280-204US
; REFERENCE/DOCKET NUMBER: DHHS Ref. No. 5652223 E-112-94/0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-212-190A-5

Query Match 44.6%; Score 25; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PLSVPC 10
Db 3 PAPVPC 8

RESULT 8
US-08-900-321-5
; Sequence 5, Application US/08900321
; Patent No. 5981712
; GENERAL INFORMATION:
; APPLICANT: Kohn, Elise C.
; APPLICANT: Liotta, Lance A.
; APPLICANT: Kim, Young S.
; TITLE OF INVENTION: DNA Encoding CAI Resistance Proteins and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,321
; FILING DATE: 25-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,190
; FILING DATE: 14-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-204100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
```

```
;
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-900-321-5

Query Match 44.6%; Score 25; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PLSVPC 10
Db 3 PAPVPC 8

RESULT 9
US-09-461-697-21
; Sequence 21, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-461-697-21

Query Match 44.6%; Score 25; DB 3; Length 10;
Best Local Similarity 44.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVP 9
Db 2 CVMTHSLP 10

RESULT 10
PCT-US95-03610-5
; Sequence 5, Application PC/TUS9503610
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA ENCODING CAI RESISTANCE PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03610
; FILING DATE: 14-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,190
```

/ FILING DATE: 14-MAR-1994  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Weber, Kenneth A.  
/ REGISTRATION NUMBER: 31,677  
/ REFERENCE/DOCKET NUMBER: 15280-204000PC  
/ REFERENCE/DOCKET NUMBER: DHS Ref. No. E-112-94/0  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (415) 543-9600  
/ TELEFAX: (415) 543-5043  
/ INFORMATION FOR SEQ ID NO: 5:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 10 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
PCT-US95-03610-5

Query Match 44.6%; Score 25; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLSVPC 10  
Db 3 PAPVPC 8

RESULT 11  
US-08-483-434A-21  
/ Sequence 21, Application US/08483434A  
/ Patent No. 5648461  
/ GENERAL INFORMATION:  
/ APPLICANT: EVAL, Jacob  
/ APPLICANT: HAMILTON, Bruce K.  
/ APPLICANT: TUSZYNSKI, George P.  
/ TITLE OF INVENTION: Synthetic Analogs of Thrombospondin and  
/ TITLE OF INVENTION: Therapeutic Use Thereof  
/ NUMBER OF SEQUENCES: 36  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
/ STREET: 1601 Market Street, 36th Floor  
/ CITY: Philadelphia  
/ STATE: Pennsylvania  
/ COUNTRY: USA  
/ ZIP: 19103-2398  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/483,434A  
/ FILING DATE: 07-JUN-1995  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/450,738  
/ FILING DATE: 25-MAY-1995  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/185,614  
/ FILING DATE: 24-JAN-1994  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/024,436  
/ FILING DATE: 01-MAR-1993  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 07/587,197  
/ FILING DATE: 24-SEP-1990  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 07/483,527  
/ FILING DATE: 22-FEB-1990  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Leary Ph.D., Kathryn  
/ REGISTRATION NUMBER: 36,317  
/ REFERENCE/DOCKET NUMBER: 9598-306 (9049)

/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (215) 567-2020  
/ TELEFAX: (215) 567-2991  
/ TELEX: 831-494  
/ INFORMATION FOR SEQ ID NO: 21:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 6 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
US-08-483-434A-21

Query Match 42.9%; Score 24; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SVPC 10  
Db 2 SVPC 5

RESULT 12  
US-08-476-134A-30  
/ Sequence 30, Application US/08476134A  
/ Patent No. 6239110  
/ GENERAL INFORMATION:  
/ APPLICANT: EVAL, JACOB  
/ APPLICANT: HAMILTON, BRUCE K.  
/ APPLICANT: TUSZYNSKI, GEORGE P.  
/ TITLE OF INVENTION: SYNTHETIC ANALOGS OF THROMBOSPONDIN AND THERAPEUTIC USE  
/ TITLE OF INVENTION: THEREOF  
/ FILE REFERENCE: 07206-0009  
/ CURRENT APPLICATION NUMBER: US/08/476,134A  
/ CURRENT FILING DATE: 1995-06-07  
/ PRIOR APPLICATION NUMBER: 07/587,197  
/ PRIOR FILING DATE: 1990-09-24  
/ PRIOR APPLICATION NUMBER: 07/483,527  
/ PRIOR FILING DATE: 1990-02-22  
/ PRIOR APPLICATION NUMBER: 08/450,738  
/ PRIOR FILING DATE: 1995-05-25  
/ PRIOR APPLICATION NUMBER: 08/185,614  
/ PRIOR FILING DATE: 1994-01-24  
/ PRIOR APPLICATION NUMBER: 08/024,436  
/ PRIOR FILING DATE: 1993-03-01  
/ NUMBER OF SEQ ID NOS: 34  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 30  
/ LENGTH: 6  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic  
/ OTHER INFORMATION: analog of thrombospondin  
US-08-476-134A-30

Query Match 42.9%; Score 24; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SVPC 10  
Db 2 SVPC 5

RESULT 13  
5190920-26  
/ Patent No. 5190920  
/ APPLICANT: EVAL, JACOB, HAMILTON, BRUCE K.; TUSZYNSKI,  
/ GEORGE P.  
/ TITLE OF INVENTION: METHOD FOR USING SYNTHETIC ANALOGS  
/ OF THROMBOSPONDIN FOR INHIBITING METASTASIS ACTIVITY  
/ NUMBER OF SEQUENCES: 32

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,197
; FILING DATE: 24-SEP-1990
; SEQ ID NO:26:
; LENGTH: 6
5190920-26

Query Match          42.9%; Score 24; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 SVPC 10
      ||||
Db      2 SVPC 5

RESULT 14
5506208-28
; Patent No. 5506208
; APPLICANT: EYAL, JACOB; HAMILTON, BRUCE K.; TUSZYNSKI,
; GEORGE P.
; TITLE OF INVENTION: METHOD FOR USING SYNTHETIC ANALOGS OF
; THROMBOSPONDIN FOR INHIBITING ANGIOGENESIS ACTIVITY
; NUMBER OF SEQUENCES: 45
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,181
; FILING DATE: 22-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 131,565
; FILING DATE: 04-OCT-1993
; APPLICATION NUMBER: 895,764
; FILING DATE: 09-JUN-1992
; APPLICATION NUMBER: 587,197
; FILING DATE: 24-SEP-1990
; SEQ ID NO:28:
; LENGTH: 6
5506208-28

Query Match          42.9%; Score 24; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 SVPC 10
      ||||
Db      2 SVPC 5

RESULT 15
US-08-331-383-35
; Sequence 35, Application US/08331383
; Patent No. 5605809
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of
; TITLE OF INVENTION: Proteases in Biological Samples and Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,383
; FILING DATE: 28-OCT-1994

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 16865-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: one-of (5)
; OTHER INFORMATION: /note= "Xaa is Met or norleucine."
; FEATURE:
; NAME/KEY: Region
; LOCATION: one-of (8)
; OTHER INFORMATION: /note= "Xaa is Pro or aminoisobutyric
; OTHER INFORMATION: acid."
; US-08-331-383-35

Query Match          42.9%; Score 24; DB 1; Length 9;
Best Local Similarity 37.5%; Pred. No. 3e+05;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      3 SVPLSVPC 10
      :|||:|
Db      2 AIPXSIXC 9

Search completed: September 5, 2004, 11:35:35
Job time : 16 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:00:28 ; Search time 11.6667 Seconds  
(without alignments)  
74.205 Million cell updates/sec

Title: US-09-761-636A-14

Perfect score: 52

Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 790

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	22	42.3	6	2	I65546
2	19	36.5	5	2	A60521
3	19	36.5	9	2	A60356
4	18	34.6	9	2	A61620
5	17	32.7	8	2	S10783
6	17	32.7	9	2	S66419
7	16	30.8	4	2	I51049
8	16	30.8	9	2	S13636
9	15	28.8	6	2	H43394
10	15	28.8	8	2	G33098
11	15	28.8	8	4	I54017
12	15	28.8	9	2	PT0080
13	14	26.9	7	2	I48105
14	14	26.9	8	2	S21288
15	14	26.9	9	2	D48186
16	14	26.9	9	2	PH0943
17	13	25.0	5	2	E42364
18	13	25.0	7	2	S42620
19	13	25.0	8	2	S39745
20	13	25.0	8	2	A42689
21	13	25.0	9	2	B28495
22	13	25.0	9	2	PT0247
23	13	25.0	9	2	PT0268
24	13	25.0	9	2	S26508
25	12	23.1	5	2	B22565
26	12	23.1	6	2	B34835
27	12	23.1	6	2	B26206
28	12	23.1	6	4	A35039
29	12	23.1	7	2	ECMUCR

30 12 23.1 7 2 E61491 seed protein ws-5  
31 12 23.1 7 2 PH1602 Ig H chain V-D-J r  
32 12 23.1 7 2 PH0932 T-cell receptor be  
33 12 23.1 8 2 S16324 hypothetical prote  
34 12 23.1 8 2 A35180 neutral proteinase  
35 12 23.1 8 2 PH0934 T-cell receptor be  
36 12 23.1 9 2 S19329 sperm-activating p  
37 12 23.1 9 2 G58502 kidney and bladder  
38 12 23.1 9 2 S55696 phosphoenolpyruvat  
39 12 23.1 9 2 AS3797 3',5'-cyclic-GMP p  
40 12 23.1 9 2 PH0935 T-cell receptor be  
41 12 23.1 9 2 PH0937 T-cell receptor be  
42 12 23.1 9 2 PH0902 T-cell receptor be  
43 12 23.1 9 2 PH0917 T-cell receptor be  
44 12 23.1 9 2 PH0918 T-cell receptor be  
45 12 23.1 9 2 PH0921 T-cell receptor be

#### ALIGNMENTS

##### RESULT 1

I65546

MHC H2-L antigen - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999

C:Accession: I65546

R:Kimura, A.; Israel, A.; Le Bail, O.; Kourilsky, P.

Cell 44, 261-272, 1986

A:Title: Detailed analysis of the mouse H-2Kb promoter: Enhancer-like sequences and their

A:Reference number: 152778; MUID:86106202; PMID:3510743

A:Accession: I65546

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6 <RES>

A:Cross-references: GB:M12483; NID:g199565; PIDN:AAA39663.1; PID:G554234

Query Match 42.3%; Score 22; DB 2; Length 6;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 LVPC 9

DB 1 MVPC 4

##### RESULT 2

A60521

glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)

N:Alternate names: glycogen phosphorylase b

C:Species: Liza ramada

C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2003

C:Accession: A60521

R:Bonamusa, L.; Baanante, I.V.

Comp. Biochem. Physiol. B 95, 295-301, 1990

A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus

A:Reference number: A60521; MUID:90227907; PMID:2109669

A:Accession: A60521

A:Molecule type: protein

A:Residues: 1-5 <BON>

C:Superfamily: glucan phosphorylase

C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein

F3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experim

Query Match 36.5%; Score 19; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISVP 5

DB 2 ISVP 5

## RESULT 3

A60356  
118k stomach cancer antigen - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 17-Mar-1999  
C;Accession: A60356  
R;Shiraishi, Y.  
Int. J. Cancer 45, 783-787, 1990  
A;Title: Western blotting analysis for malignant lymphoma and stomach cancer antigens fr  
A;Reference number: A60356; MUID:90216080; PMID:2323853  
A;Accession: A60356  
A;Molecule type: protein  
A;Residues: 1-9 <SHI>  
C;Keywords: glycoprotein

Query Match 36.5%; Score 19; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPLVP 8  
|:|:|  
Db 1 IPLKP 5

## RESULT 4

A61620  
locustamytotropin III - migratory locust  
C;Species: Locusta migratoria (migratory locust)  
C;Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 11-Jul-1997  
C;Accession: A61620  
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; Kochansky, J.P.; De Loof, A.  
Insect Biochem. Mol. Biol. 22, 447-452, 1992  
A;Title: Isolation, identification and synthesis of locustamytotropin III and IV, two add  
A;Reference number: A61620  
A;Accession: A61620  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-9 <SCH>  
C;Keywords: amidated carboxyl end; neuropeptide  
F;9/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 34.6%; Score 18; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLVP 8  
|:|:|  
Db 4 PFVP 7

## RESULT 5

S10783  
enamelin f - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 19-Mar-1997 #sequence\_revision 21-Nov-1998 #text\_change 21-Nov-1998  
C;Accession: S10783  
R;Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu  
A;Reference number: S10780; MUID:90336641; PMID:2379503  
A;Accession: S10783  
A;Molecule type: protein  
A;Residues: 1-8 <STR>  
C;Keywords: enamel; phosphoprotein

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPLVP 8  
|:|:|  
Db 1 MPLPP 5

## RESULT 6

S66419  
tetrameric proteinase precursor, 60K, dithiothreitol-sensitive - spinach (fragment)  
C;Species: Spinacia oleracea (spinach)  
C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
C;Accession: S66419  
R;Kuwabara, T.  
FEBS Lett. 371, 195-198, 1995  
A;Title: The 60-kDa precursor to the dithiothreitol-sensitive tetrameric protease of spi  
A;Reference number: S66419; MUID:95402209; PMID:7672127  
A;Accession: S66419  
A;Molecule type: protein  
A;Residues: 1-9 <KW>

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLVP 8  
|:|:|  
Db 2 PILP 5

## RESULT 7

I51049  
metallothionein-A - rainbow trout (fragment)  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C;Accession: I51049  
R;Olsson, P.E.; Kling, P.; Erkel, L.J.; Kille, P.  
Eur. J. Biochem. 230, 344-349, 1995  
A;Title: Structural and functional analysis of the rainbow trout (Oncorhynchus mykiss) me  
A;Reference number: I51049; MUID:95324545; PMID:7601121  
A;Accession: I51049  
A;Status: preliminary; translated from GB/EMBL/DDBU  
A;Molecule type: DNA  
A;Residues: 1-4 <OLS>  
A;Cross-references: EMBL:X60181; NID:g1019799; PIDN:CAA56466.1; PID:g4379328

Query Match 30.8%; Score 16; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PC 9  
|:|  
Db 3 PC 4

## RESULT 8

S13636  
coat protein beta chain, Golgi-derived - rabbit (fragment)  
N;Alternate names: beta-COP protein  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 19-Mar-1997 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C;Accession: S13636  
R;Serafini, T.; Stenbeck, G.; Brecht, A.; Lottspeich, F.; Orzi, L.; Rothman, J.E.; Wiel  
Nature 349, 215-220, 1991  
A;Title: A coat subunit of Golgi-derived non-clathrin-coated vesicles with homology to  
A;Reference number: S13636; MUID:91101693; PMID:1898984  
A;Accession: S13636  
A;Molecule type: protein  
A;Residues: 1-9 <SER>  
C;Superfamily: coatomer complex beta chain  
C;Keywords: Golgi apparatus; protein transport

Query Match 30.8%; Score 16; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPLV 7  
|:|:|  
Db 5 IPIV 8



```

RESULT 9
H48394
glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr
C:Species: Bos primigenius taurus (cattle)
C>Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997
C:Accession: H48394
R:Mathier, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A:Reference number: A48394; MUID:93250576; PMID:8485470
A:Accession: H48394
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-6 <MAT>
A:Experimental source: milk
A>Note: sequence extracted from NCBI backbone (NCBIP:131518)
C:Keywords: glycoprotein

Query Match      28.8%; Score 15; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches      3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      4 VPLVPC 9
      |||
Db      1 VELLGC 6

RESULT 10
G33098
205K exoantigen - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C>Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
C:Accession: G33098
R:Nichols, J.H.; Hager, L.P.
submitted to the Protein Sequence Database, May 1990
A:Reference number: A33098
A:Accession: G33098
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <NIC>

Query Match      28.8%; Score 15; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches      3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 VPL 6
      |||
Db      2 VPL 4

RESULT 11
I54017
granulocyte-colony stimulating factor precursor - synthetic (fragment)
C:Species: synthetic
A>Note: human gene engineered and expressed in Escherichia coli
C>Date: 28-Jan-2000 #sequence_revision 28-Jan-2000 #text_change 28-Jan-2000
C:Accession: I54017
R:Devlin, P.B.; Drummond, R.J.; Toy, P.; Mark, D.F.; Watt, K.W.; Devlin, J.J.
Gene 65, 13-22, 1988
A:Title: Alteration of amino-terminal codons of human granulocyte-colony-stimulating fac
i.
A:Reference number: I54017; MUID:88284374; PMID:2456256
A:Accession: I54017
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-8 <DEV>
A:Cross-references: GB:M20922; NID:9806638; PIDN:AAAG6353.1; PID:gl83043

Query Match      28.8%; Score 15; DB 4; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;

```

```

Matches      3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 PLVP 8
      |||
Db      3 PLGP 6

RESULT 12
PT0080
60K Ca binding protein - edible frog (fragment)
C:Species: Rana esculenta (edible frog)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: PT0080
R:Trevoso, S.; Zorzato, F.; Chiozzi, P.; Melandri, P.; Volpe, P.; Pozzan, T.
Biochem. Biophys. Res. Commun. 175, 444-450, 1991
A:Title: Frog brain expresses a 60 kDa Ca2+ binding protein similar to mammalian calret;
A:Reference number: PT0080; MUID:91207333; PMID:2018493
A:Accession: PT0080
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <IRE>

Query Match      28.8%; Score 15; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches      3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 PLV 7
      |||
Db      2 PLV 4

RESULT 13
I48105
dihydrofolate reductase - Chinese hamster (fragment)
C:Species: Cricetus griseus (Chinese hamster)
C>Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999
C:Accession: I48105
R:Azizkhan, J.C.; Vaughn, J.P.; Christy, R.J.; Hamlin, J.L.
Biochemistry 25, 6228-6236, 1986
A:Title: Nucleotide sequence and nuclease hypersensitivity of the Chinese hamster dihydr
A:Reference number: I48105; MUID:87076541; PMID:3024702
A:Accession: I48105
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-7 <RES>
A:Cross-references: GB:M14771; NID:gl91055; PIDN:AAA36975.1; PID:gl91056

Query Match      26.9%; Score 14; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches      2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 PLV 7
      |||
Db      4 PLI 6

RESULT 14
S21288
lectin - potato (fragment)
C:Species: Solanum tuberosum (potato)
C>Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 05-Dec-1998
C:Accession: S21288
R:Millar, D.J.; Allen, A.K.; Smith, C.G.; Sidebottom, C.; Slabas, A.R.; Bolwell, G.P.
Biochem. J. 283, 813-821, 1992
A:Title: Chitin-binding proteins in potato (Solanum tuberosum L.) tuber. Characterization
A:Reference number: S21288; MUID:92272683; PMID:1590771
A:Accession: S21288
A:Molecule type: protein
A:Residues: 1-8 <MIL>
A:Experimental source: var. Ulster Sceptre
C:Function:
A:Description: may be involved in defence mechanism of the plant
C:Keywords: hydroxyproline; lectin

```

Query Match  
Best Local Similarity 26.9%; Score 14; DB 2; Length 8;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SVPLVP 8  
| | |  
Db 3 STFSPP 8

RESULT 15  
D48186  
AtPase R1 subunit - wood tobacco (fragment)  
C/Species: Nicotiana sylvestris (wood tobacco)  
C/Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Feb-1997  
C/Accession: D48186  
R/De Paeppe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993  
A/Title: Specific mitochondrial proteins in pollen: presence of an additional ATP synthase  
A/Reference number: A48186; MUID:93317598; PMID:8327463  
A/Accession: D48186  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-9 <DE1>  
A/Experimental source: pollen  
A/Note: sequence extracted from NCBI backbone (NCBIP:134871)

Query Match  
Best Local Similarity 60.0%; Score 14; DB 2; Length 9;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPLVP 8  
| | |  
Db 3 VDLAP 7

Search completed: September 5, 2004, 11:06:22  
Job time : 11.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 10:57:02 ; Search time 7 Seconds  
(without alignments)  
66.947 Million cell updates/sec

Title: US-09-761-636A-14  
Perfect score: 52  
Sequence: 1 C1SVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	38.5	9	1	MGMT_BOVIN	P29177 bos taurus
2	18	34.6	8	1	COM2_CONPU	P58785 conus purpur
3	18	34.6	9	1	LMT3_LOCFI	P41489 locusta mig
4	17	32.7	9	1	UPA3_HUMAN	P30089 homo sapien
5	16	30.8	8	1	PKX2_PERAM	P82692 periplaneta
6	15	28.8	8	1	PKX2_PERAM	P82692 periplaneta
7	15	28.8	9	1	UPA7_HUMAN	P30093 homo sapien
8	13	25.0	6	1	E101_LITRU	P82096 litoria rub
9	13	25.0	9	1	CONO_CONST	P05487 conus stria
10	12	23.1	7	1	CARP_MYTED	P10420 mytilus edu
11	12	23.1	8	1	AL16_CVDPO	P82157 cydia pomon
12	12	23.1	8	1	FUSS_FUSSO	P81010 fusarium so
13	12	23.1	8	1	UPA1_HUMAN	P30087 homo sapien
14	12	23.1	9	1	FAR5_PENMO	P83320 penaeus mon
15	12	23.1	9	1	FLA2_TREHY	P80159 treponema h
16	12	23.1	9	1	OXYT_BUFRE	P42995 bufo regula
17	12	23.1	9	1	SAP_STOVA	P24047 stomopneute
18	11	21.2	6	1	VP19_HSVIK	P23210 herpes simp
19	11	21.2	7	1	CCR1_ENTFA	P20104 enterococcu
20	11	21.2	7	1	MNP1_LEPDE	P42984 leptonotars
21	11	21.2	7	1	TPFY_FACDA	P83455 pachymedusa
22	11	21.2	8	1	VGIG_HSV2B	P81780 herpes simp
23	11	21.2	9	1	COXE_THUOB	P80975 thunnus obe
24	11	21.2	9	1	DNF1_LOCFI	P16339 locusta mig
25	11	21.2	9	1	FAR9_ASCSU	P43172 ascaris suu
26	11	21.2	9	1	OXYA_SQUAC	P42999 squalus aca
27	11	21.2	9	1	OXYT_RABIT	P32878 oryctolagus
28	11	21.2	9	1	RT33_BOVIN	P82926 bos taurus
29	10	19.2	7	1	UF51_LITRU	P82065 litoria rub
30	10	19.2	7	1	UF04_MOUSE	P38642 mus musculus
31	10	19.2	9	1	COW CONVE	P83047 conus ventr
32	10	19.2	9	1	PKX1_PERAM	P82691 periplaneta
33	10	19.2	9	1	TAL1_PICJA	P17440 pichia jadi

```

34      10      19.2      9      1      TAL3_PICJA
35      9      17.3      5      1      E103_LITRU
36      9      17.3      5      1      E104_LITRU
37      9      17.3      6      1      C1P1_MYTED
38      9      17.3      6      1      C1P2_MYTED
39      9      17.3      7      1      BRHP_CONIM
40      9      17.3      7      1      UN06_PINPS
41      9      17.3      7      1      WWA1_ACHFU
42      9      17.3      7      1      WWA2_ACHFU
43      9      17.3      7      1      WWA3_ACHFU
44      9      17.3      8      1      ACT_CARMA
45      9      17.3      8      1      AKH_LTBAD

```

```

P17441 pichia jadi
P82099 litoria rub
P82100 litoria rub
P13736 mytilus edu
P13737 mytilus edu
P58803 conus imper
P81675 pinus pinas
P35919 achatina fu
P35920 achatina fu
P35921 achatina fu
P80709 carcinus ma
P25418 libellula a

```

## ALIGNMENTS

```

RESULT 1
MGMT_BOVIN          STANDARD;          PRT;          9 AA.
AC      P29177;
DT      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63) (6-O-
DE      methylguanine-DNA methyltransferase) (Fragment).
GN      MGMT.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RN      [1]__TaxID=9913;
RP      SEQUENCE.
RC      TISSUE=Thymus;
RX      MEDLINE=90174912; PubMed=2308822;
RA      Rydberg B., Hall J., Karran P.;
RT      "Active site amino acid sequence of the bovine O6-methylguanine-DNA
RT      methyltransferase.";
RL      Nucleic Acids Res. 18:17-21(1990).
CC      -!- FUNCTION: Repair of alkylated guanine in DNA by stoichiometrically
CC      transferring the alkyl group at the O-6 position to a cysteine
CC      residue in the enzyme. This is a suicide reaction: the enzyme is
CC      irreversibly inactivated.
CC      -!- CATALYTIC ACTIVITY: DNA (containing 6-O-methylguanine) +
CC      [protein]-L-cysteine = DNA (without 6-O-methylguanine) + protein
CC      S-methyl-L-cysteine.
CC      -!- SIMILARITY: WITH SEGMENTS OF E.COLI ADA AND OGT METHYLTRANSFERASE
CC      WHICH ENCOMPASS THE ALKYL-ACCEPTOR RESIDUES.
DR      InterPro: IPR001497; Methyltransf_1.
DR      PROSITE: PS00374; MGMT; PARTIAL.
KW      DNA repair; transferase; Methyltransferase.
FT      NON_TER      1      1
FT      ACT_SITE      9      9      ALKYL GROUP ACCEPTOR (BY SIMILARITY).
FT      NON_TER      9      9
SQ      SEQUENCE      9 AA;      967 MW;      325171A720476047 CRC64;

```

Query Match 38.5%; Score 20; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.4e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      6 LVPC 9
Db      6 LTTC 9

```

```

RESULT 2
COW2_CONPU          STANDARD;          PRT;          8 AA.
AC      P58785;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Leu-contryphan-P.
OS Conus purpurascens (Purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=41690;
RN [1]_SEQUENCE;
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC STRAIN=Clipperton Island; TISSUE=Venom;
RX MEDLINE=9938839; PubMed=10461743;
RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,
RA Olivera B.M.;
RT "A novel D-leucine-containing Conus peptide: diverse conformational
RT dynamics in the contryphan family.";
RL J. Pept. Res. 54:93-99(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.
CC -1- SIMILARITY: Belongs to the contryphan family.
KW Toxin; Hydroxylation; D-amino acid.
FT DISULFID 2 8
FT MOD RES 4 4 D-LEUCINE.
SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;

Query Match 34.6%; Score 18; DB 1; Length 8;
Best Local Similarity 40.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISVP 5
| : |
| : |
DB 2 CVLLP 6

RESULT 3
LMT3 LOCM1
ID LMT3 LOCM1 STANDARD; PRT; 9 AA.
AC P41489;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Locustamytropin 3 (LOW-MT-3).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]_SEQUENCE, AND SYNTHESIS.
RP TISSUE=Brain;
RC Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,
RA de Loof A.;
RT "Isolation, identification and synthesis of locustamytropin III and
RT IV, two additional neuropeptides of Locusta migratoria: members of the
RT locustamytropin peptide family.";
RL Insect Biochem. Mol. Biol. 22:447-452(1992).
CC -1- FUNCTION: Potent mediator of visceral muscle contractile activity
CC (myotropic activity).
CC -1- SIMILARITY: Belongs to the pyrokinin family.
DR PIR; A61620;
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; 1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1140 MW; D5AE1772C9D776C6 CRC64;

Query Match 34.6%; Score 18; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLVP 8
| | |
| | |
DB 4 PFVP 7

```

```

RESULT 4
UPA3 HUMAN STANDARD; PRT; 9 AA.
AC P30089;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of plasma (Spot 11) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE.
RP TISSUE=Plasma;
RC MEDLINE=93092937; PubMed=1459097;
RX Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 4.6, its MW is: 46 kDa.
DR SWISS-2DPAGE; P30089; HUMAN.
FT NON_TER 1 1
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1056 MW; 26F2B1BAF769C737 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLVP 8
| | |
| | |
DB 2 PLFP 5

RESULT 5
PPK2 PERAM STANDARD; PRT; 8 AA.
AC P82692;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyrokinin-2 (Pea-PK-2) (EXPRL-amide).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Hexapoda; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]_SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RP TISSUE=Corpora cardiaca;
RC MEDLINE=97353923; PubMed=9210163;
RX Predel R., Kellner R., Kaufmann R., Penzlin H., Gaede G.;
RA "Isolation and structural elucidation of two pyrokinins from the
RT retrocerebral complex of the American cockroach.";
RL Peptides 18:473-478(1997).
RN [2]_TISSUE SPECIFICITY.
RP TISSUE SPECIFICITY.
RX Predel R., Eckert M.;
RA "Tagma-specific distribution of FXPRLAMides in the nervous system of
RT the American cockroach.";
RL J. Comp. Neurol. 419:352-363(2000).
CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic
CC activity).
CC -1- TISSUE SPECIFICITY: Corpora cardiaca.
CC -1- MASS SPECTROMETRY: MW=883; METHOD=MALDI.
CC -1- SIMILARITY: Belongs to the pyrokinin family.
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; FALSE_NEG.

```

```

KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 884 MW; C834176DD9D7775 CRC64;

Query Match 30.8%; Score 16; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SVPLVP 8
DB 1 SPFFAP 6

RESULT 6
PPK3 PERAM
ID _PPK3 PERAM STANDARD; PRT; 8 AA.
AC P826T8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyrokinin-3 (Pea-PK-3) (FXPRL-amide).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattelloidea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-Retrocerebral complex;
RX MEDLINE=59212469; PubMed=10196736;
RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;
RT "Differential distribution of pyrokinin-isoforms in cerebral and
RT abdominal neurohemal organs of the American cockroach.";
RL Insect Biochem. Mol. Biol. 29:139-144 (1999).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
RL J. Comp. Neurol. 419:352-363 (2000).
CC -!- FUNCTION: Mediates visceral muscle contractile activity (myotropic
CC activity).
CC -!- TISSUE SPECIFICITY: Corpora cardiaca.
CC -!- MASS SPECTROMETRY: MW=956.5; METHOD=WALDI.
CC -!- SIMILARITY: Belongs to the pyrokinin family.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 997 MW; 0B34177409D772C7 CRC64;

Query Match 28.8%; Score 15; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LVP 8
DB 1 LVP 3

RESULT 7
UPA7 HUMAN
ID _UPA7 HUMAN STANDARD; PRT; 9 AA.
AC P30053;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of plasma (Spot 18) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.

TISSUE=Plasma;
RC MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714 (1992).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 5.05, its MW is: 37 kDa.
DR SWISS-2DPAGE; P30093; HUMAN.
FT NON_TER 1 1
FT UNSURE 5 5
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1042 MW; 5C14477AEB0772C7 CRC64;

Query Match 28.8%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LVP 8
DB 2 LVP 4

RESULT 8
EI01 LITRU
ID _EI01 LITRU STANDARD; PRT; 6 AA.
AC P82096;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645 (1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPL 6
DB 2 VPI 4

RESULT 9
CONO CONST
ID _CONO CONST STANDARD; PRT; 9 AA.
AC P05487;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arg-conopressin S.
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbecoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;

```



```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 4.9, its MW is: 65 kDa.
CX SWISS-2DPAGE; P30087; HUMAN.
FT NON_TER 1
FT UNSURE 8
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 23.1%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVP 5
Db 5 NVP 7

RESULT 14
FAR5 PENMO
ID FAR5 PENMO STANDARD; PRT; 9 AA.
AC P83320;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE FMRamide-like neuropeptide FLPS (SMPSLRFR-amide).
OS Penaeus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OX NCBI_TaxID=6687;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Eyestalk;
RX MEDLINE=21956277; PubMed=11959015;
RA Sithigorngul P., Pupum J., Krungkarn C., Longyant S.,
RA Chaivithangkura P., Sithigorngul W., Petsom A.;
RT "Seven novel FMRamide-like neuropeptide sequences from the eyestalk
RT of the giant tiger prawn Penaeus monodon.";
RL Comp. Biochem. Physiol. 131B:325-337(2002).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=1121.1; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
CC family.
DR GO:0007218; P:neuropeptide signaling pathway; TAS.
KW Neuropeptide; Amidation.
FT MOD_RES 9
FT MOD_RES 9
SQ SEQUENCE 9 AA; 1106 MW; B60B07340735A766 CRC64;

Query Match 23.1%; Score 12; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVP 5
Db 1 SMP 3

RESULT 15
FLA2 TREHY
ID FLA2 TREHY STANDARD; PRT; 9 AA.

```

```

AC P80159;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar filament outer layer protein flaA2 (35 kDa sheath protein)
DE (Fragment).
GN FLA2
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
OX NCBI_TaxID=159;
RN [1]
RP SEQUENCE.
RC STRAIN=C5;
RX MEDLINE=93139764; PubMed=1487733;
RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA van der Zeijst B.A.M., Kusters J.G.;
RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT composed of two sheath proteins and three core proteins.";
RL J. Gen. Microbiol. 138:2697-2706(1992).
CC -!- FUNCTION: Component of the outer layer of the flagella.
CC -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
CC SHEATH PROTEINS, FLA1 (44 kDa) AND FLA2 (35 kDa) AROUND A CORE
CC THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND
CC FLAB3 (32 kDa).
CC -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
KW Flagellum; Periplasmic.
FT UNSURE 2
FT UNSURE 8
FT NON_TER 9
SQ SEQUENCE 9 AA; 1129 MW; 855A19C68B4772D1 CRC64;

Query Match 23.1%; Score 12; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVP 5
Db 2 TVP 4

Search completed: September 5, 2004, 11:04:00
Job time : 7 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 10:59:47 ; Search time 31.6667 Seconds  
(without alignments)  
89.674 Million cell updates/sec

Title: US-09-761-636A-14  
Perfect score: 52  
Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 840

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_xvirus.\*  
16: sp\_bacteriaph.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	38.5	9	3 Q9P8E5	Q9P8E5 kluyveromyc
2	19	36.5	7	13 Q42564	Q42564 figu rubrip
3	19	36.5	8	2 Q56140	Q56140 streptococ
4	18	34.6	9	13 Q8AYL5	Q8AYL5 carassius a
5	18	34.6	9	13 Q8AUM7	Q8AUM7 carassius a
6	17	32.7	9	10 Q7X6A3	Q7X6A3 zea mays su
7	16	30.8	7	10 P93233	P93233 lycopersico
8	16	30.8	8	6 O02831	O02831 oryctolagus
9	16	30.8	8	6 Q9TRY3	Q9TRY3 sus sp. ins
10	16	30.8	9	4 Q16220	Q16220 homo sapien
11	16	30.8	9	4 Q9UMF3	Q9UMF3 homo sapien
12	16	30.8	9	12 Q8QVD3	Q8QVD3 ovine respi
13	15	28.8	7	15 Q07624	Q07624 rous sarcom
14	15	28.8	9	7 O78225	O78225 mus musculu
15	15	28.8	9	10 Q9S8J8	Q9S8J8 oryza sativ
16	15	28.8	9	11 Q35953	Q35953 mus musculu

17	14	26.9	7	2 Q07354	Q07354 synchococc
18	14	26.9	8	2 Q9X3K1	Q9X3K1 prochloroco
19	14	26.9	8	2 P83532	P83532 lactobacill
20	14	26.9	8	6 Q9TRX8	Q9TRX8 bos taurus
21	13	25.0	6	5 P83569	P83569 sepiia offic
22	13	25.0	7	12 Q67113	Q67113 influenzavi
23	13	25.0	8	2 Q49534	Q49534 mycoplasma
24	13	25.0	8	2 Q32560	Q32560 escherichia
25	13	25.0	8	8 Q8WFR5	Q8WFR5 diadema pau
26	13	25.0	8	11 Q8K3Z7	Q8K3Z7 mus musculu
27	13	25.0	8	11 Q99P40	Q99P40 mus musculu
28	13	25.0	8	11 P82598	P82598 rattus norv
29	13	25.0	8	12 Q9E8Q5	Q9E8Q5 beest soil-b
30	13	25.0	8	12 Q9E8Q2	Q9E8Q2 beest soil-b
31	13	25.0	8	12 Q9E8P9	Q9E8P9 beest soil-b
32	13	25.0	8	12 Q9E8Q3	Q9E8Q3 beest soil-b
33	13	25.0	8	12 Q9E8Q1	Q9E8Q1 beest soil-b
34	13	25.0	8	12 Q9DSN2	Q9DSN2 beest soil-b
35	13	25.0	8	12 Q9E8Q4	Q9E8Q4 beest soil-b
36	13	25.0	8	12 Q9DSN5	Q9DSN5 beest soil-b
37	13	25.0	8	12 Q9E8P7	Q9E8P7 beest soil-b
38	13	25.0	8	12 Q9DSN1	Q9DSN1 beest soil-b
39	13	25.0	8	12 Q9DSN3	Q9DSN3 beest soil-b
40	13	25.0	8	12 Q9DSN4	Q9DSN4 beest soil-b
41	13	25.0	8	12 Q9E8Q0	Q9E8Q0 beest soil-b
42	13	25.0	8	12 Q9E8Q7	Q9E8Q7 beest soil-b
43	13	25.0	8	12 Q9DSN6	Q9DSN6 beest soil-b
44	13	25.0	8	12 Q9DSN0	Q9DSN0 beest soil-b
45	13	25.0	8	12 Q9E8Q6	Q9E8Q6 beest soil-b

#### ALIGNMENTS

#### RESULT 1

Q9P8E5 PRELIMINARY; PRT; 9 AA.  
AC Q9P8E5, AC Q9P8E5, (TREMREL. 15, Created)  
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)  
DE HIS4 protein (Fragment).  
GN HIS4.  
OS Kluyveromyces lactis (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
OX NCBI\_TaxID=28985;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRRL-Y1140;  
RX MEDLINE=99448382; PubMed=10518937;  
RA Lamas-Maceiras M., Esperanza Cerdan E., Freire-Picos M.A.;  
RT "Kluyveromyces lactis HIS4 transcriptional regulation: similarities  
RT and differences to Saccharomyces cerevisiae HIS4 gene.";  
RL FEBS Lett. 458:72-76(1999).  
DR EMBL; AJ238494; CAB87125.1; .  
FT NON\_TER  
SQ SEQUENCE 9 AA; 1015 MW; 5770D2D772D2D767 CRC64;

Query Match 38.5%; Score 20; DB 3; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPLVP 8  
Db 2 LPVVP 6

#### RESULT 2

O42564 PRELIMINARY; PRT; 7 AA.  
ID O42564  
AC O42564;  
DT 01-JAN-1998 (TREMREL. 05, Created)

DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Truncated voltage-gated sodium channel alpha subunit (Fragment).  
 GN SCN8A.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97442476; PubMed=9295353;  
 RA Plummer N.W., McBurney M.W., Meisler M.H.;  
 RT "Alternative splicing of the sodium channel SCN8A predicts a truncated  
 RT two-domain protein in fetal brain and non-neuronal cells.";  
 RL J. Biol. Chem. 272:24008-24015(1997).  
 DR EMBL; U97673; AAB80916.1; -.  
 DR GO; GO:0005216; F:ion channel activity; IEA.  
 KW Ionic channel.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 730 MW; 75872EA2C73772A0 CRC64;  
 Query Match 36.5%; Score 19; DB 13; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 VPLV 7  
 Db ||||  
 1 VPLV 4  
 RESULT 3  
 ID Q56140 PRELIMINARY; PRT; 8 AA.  
 AC Q56140;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE STP6 protein (Fragment).  
 GN STP6.  
 OS Streptococcus thermophilus.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1308;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ST11;  
 RX MEDLINE=95047254; PubMed=7958782;  
 RA Constable A., Mollet B.;  
 RT "Isolation and characterization of promoter regions from Streptococcus  
 RT thermophilus.";  
 RL FEMS Microbiol. Lett. 122:85-90(1994).  
 DR EMBL; X78210; CAA55045.1; -.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 846 MW; ED086772D5B045B6 CRC64;  
 Query Match 36.5%; Score 19; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ISVP 5  
 Db ||||  
 3 ISVP 6  
 RESULT 4  
 ID Q8AYL5 PRELIMINARY; PRT; 9 AA.  
 AC Q8AYL5;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Cytochrome P450 aromatase (Fragment).  
 GN CYP19A.  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;  
 RT "Promoter characteristics of two CYP19 genes differentially expressed  
 RT in the brain and ovary of teleost fish.";  
 RL J. Steroid Biochem. Mol. Biol. 0:0-0(2001).  
 DR EMBL; AF324897; AAN32618.1; -.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1060 MW; C49E76D7272B040D CRC64;  
 Query Match 34.6%; Score 18; DB 13; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1e+06;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 LVPC 9  
 Db ||||  
 6 LVPC 9  
 RESULT 5  
 ID Q8AUM7 PRELIMINARY; PRT; 9 AA.  
 AC Q8AUM7;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Cytochrome P450 aromatase (Fragment).  
 GN CYP19A.  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;  
 RT "Promoter characteristics of two CYP19 genes differentially expressed  
 RT in the brain and ovary of teleost fish.";  
 RL J. Steroid Biochem. Mol. Biol. 0:0-0(2001).  
 DR EMBL; AF324895; AAN32616.1; -.  
 DR EMBL; AF324896; AAN32617.1; -.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 961 MW; C49E76D7272B187D CRC64;  
 Query Match 34.6%; Score 18; DB 13; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1e+06;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 LVPC 9  
 Db ||||  
 6 LVPC 9  
 RESULT 6  
 ID Q7X6A3 PRELIMINARY; PRT; 9 AA.  
 AC Q7X6A3;  
 DT 01-OCT-2003 (TReMBLrel. 25, Created)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Isoamylase (Fragment).  
 GN SUI.  
 OS Zea mays subsp. mays (maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.

```

OX NCBI_TaxID=4578;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. 38-11, and cv. A632;
RA Wilson L.M., Whitt S.R., Ibanez A.M., Rocheford T.R., Goodman M.M.,
RA Buckler E.S. IV.;
RT "Dissection of maize starch production by candidate gene
RT association."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY290305; AAP45331.1; -.
DR EMBL; AY290311; AAP45337.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 976 MW; DF9BCEA76736C6DD CRC64;

Query Match 32.7%; Score 17; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 VPC 9
Db :|||
5 LPC 7

RESULT 7
P93233
ID P93233 PRELIMINARY; PRT; 7 AA.
AC P93233;
DT 01-MAY-1997 (T-EMBLrel. 03, Created)
DT 01-MAY-1997 (T-EMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
DE (Fragment).
GN LE-ACS1B.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97351561; PubMed=9207843;
RA Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate
RT synthase genes by elicitor in suspension cultures of tomato
RT (Lycopersicon esculentum).";
RL Plant Mol. Biol. 34:275-286 (1997).
DR EMBL; U75692; AAC49682.1; -.
DR GO; GO:0016847; F1-aminocyclopropane-1-carboxylate synthase . . .; IEA.
DR GO; GO:0016829; F1-yase activity; IEA.
KW Lyase.
FT NON_TER 1
SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 30.8%; Score 16; DB 10; Length 7;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLV 7
Db :|||
1 SRPLV 5

RESULT 8
O02831
ID O02831 PRELIMINARY; PRT; 8 AA.
AC O02831;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Pro alpha 1 type III collagen protein (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

```

```

OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96377339; PubMed=8783186;
RA Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA Vuorio E.;
RT "Evidence for insufficient chondrocytic differentiation during repair
RT of full-thickness defects of articular cartilage."
RL Matrix Biol. 15:39-47 (1996).
DR EMBL; S83371; AAD14433.1; -.
KW Collagen.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

Query Match 30.8%; Score 16; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PC 9
Db :|||
3 PC 4

RESULT 9
Q9TRY3
ID Q9TRY3 PRELIMINARY; PRT; 8 AA.
AC Q9TRY3;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Insulin-like growth factor-binding protein-6, IGFBP-6 (Fragment).
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]
RP SEQUENCE.
RX MEDLINE=92049376; PubMed=1719383;
RA Shimasaki S., Gao L., Shimomura M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
RT protein-6."
RL Mol. Endocrinol. 5:938-948 (1991).
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 850 MW; 9FB2CEA37EA7687D CRC64;

Query Match 30.8%; Score 16; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PC 9
Db :|||
3 PC 4

RESULT 10
Q16220
ID Q16220 PRELIMINARY; PRT; 9 AA.
AC Q16220;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE HGRP protein (Fragment).
GN HGRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94320083; PubMed=8044796;
RA Nagalla S.R., Spindel E.R.;
RT "Functional analysis of the 5'-flanking region of the human gastrin-

```

```

RT   releasing peptide gene in small cell lung carcinoma cell lines.";
RL   Cancer Res. 54:4461-4467(1994).
DR   EMBL; S73265; AAD14116.1; -.
DR   GO; GO:0005634; C:nucleus; NAS.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
FT   NON TER          9
SQ   SEQUENCE          9 AA; 1070 MW; 77FE37672B040864 CRC64;

Query Match          30.8%; Score 16; DB 4; Length 9;
Best Local Similarity 75.0%; Pred. No. 1e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY   4 VPLV 7
DB   :|||
      6 LPLV 9

RESULT 11
Q9UMF3 PRELIMINARY; PRT; 9 AA.
ID   Q9UMF3
AC   Q9UMF3
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE   PD-1 protein (Fragment).
GN   PD-1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=97473511; PubMed=9332365;
RA   Finger L.R., Pu J., Wasserman R., Vibhakar R., Louie E., Hardy R.R.,
RA   Burrows P.D., Billips L.G.;
RT   "The human PD-1 gene: complete cDNA, genomic organization, and
RT   developmentally regulated expression in B cell progenitors.";
RL   Gene 197:177-187(1997).
DR   EMBL; U64864; AAC51774.1; -.
FT   NON TER          9
SQ   SEQUENCE          9 AA; 1067 MW; DD4A676DC6C76046 CRC64;

Query Match          30.8%; Score 16; DB 4; Length 9;
Best Local Similarity 28.6%; Pred. No. 1e+06;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY   2 ISVPLV 8
DB   : :|
      1 MQIQAP 7

RESULT 12
Q8QVD3 PRELIMINARY; PRT; 9 AA.
ID   Q8QVD3
AC   Q8QVD3
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE   Matrix protein 2 (fragment).
GN   M2.
OS   Ovine respiratory syncytial virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX   NCBI_TaxID=28869;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21580659; PubMed=11724134;
RA   Eleraky N.Z., Kania S.A., Potgieter L.N.;
RT   "The ovine respiratory syncytial virus F gene sequence and its
RT   diagnostic application.";
RL   J. Vet. Diagn. Invest. 13:455-461 (2001).
DR   EMBL; AF334398; AAL91343.1; -.

```

```

FT   NON TER          9
SQ   SEQUENCE          9 AA; 1154 MW; 8B6A3EA764541415 CRC64;

Query Match          30.8%; Score 16; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   8 PC 9
DB   |||
      6 PC 7

RESULT 13
Q07624 PRELIMINARY; PRT; 7 AA.
ID   Q07624
AC   Q07624
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE   UORF1.
DE   Rous sarcoma virus (strain Prague C).
OS   Rous sarcoma virus; Retroviridae; Alpharetrovirus.
OC   Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX   NCBI_TaxID=11888;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93010967; PubMed=1327749;
RA   Donze O., Spahr P.F.;
RA   "Role of the open reading frames of Rous sarcoma virus leader RNA in
RT   translation and genome packaging.";
RL   EMBO J. 11:3747-3757(1992).
DR   EMBL; X67587; CAA47862.1; -.
FT   NON TER          7
SQ   SEQUENCE          7 AA; 672 MW; 776045A7687DD6F0 CRC64;

Query Match          28.8%; Score 15; DB 15; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY   5 PLVP 8
DB   ||:|
      4 PSIP 7

RESULT 14
O78225 PRELIMINARY; PRT; 9 AA.
ID   O78225
AC   O78225
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE   Lymphocyte antigen (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=CRO435; TISSUE=Kidney;
RX   MEDLINE=88084418; PubMed=3692165;
RA   Golubic M., Budimir O., Schoepfer R., Kasahara M., Mayer W.E.,
RA   Figueroa F., Klein J.;
RT   "Nucleotide sequence analysis of class II genes borne by mouse t
RT   chromosomes.";
RL   Genet. Res. 50:137-146(1987).
DR   EMBL; L38589; AAA57293.1; -.
FT   NON TER          1
FT   NON TER          9
FT   NON TER          9
SQ   SEQUENCE          9 AA; 1143 MW; 9E681772C729C33A CRC64;

Query Match          28.8%; Score 15; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   6 LVP 8

```

```
Db          5 LVP 7
||||
RESULT 15
Q9S8J8
ID Q9S8J8      PRELIMINARY;      PRT;      9 AA.
AC Q9S8J8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE OryzATENSIN=BIOACTIVE peptide.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE.
RX MEDLINE=95102521; PubMed=7804141;
RA Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;
RT "Isolation and characterization of oryzatensin: a novel bioactive
RT peptide with ileum-contracting and immunomodulating activities derived
RT from rice albumin.";
RL Biochem. Mol. Biol. Int. 33:1151-1158(1994).
DR Gramene; Q9S8J8; -.
SQ SEQUENCE 9 AA; 1093 MW; 0E8C67377B56877B CRC64;

Query Match      28.8%; Score 15; DB 10; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          5 PLVP 8
          | : |
Db          3 PMYP 6

Search completed: September 5, 2004, 11:05:43
Job time : 32.6667 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 10:56:37 ; Search time 37.3333 Seconds  
(without alignments)  
68.114 Million cell updates/sec

Title: US-09-761-636A-14  
Perfect score: 52  
Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 231240

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	9	4	AAU04533
2	27	51.9	9	5	ABG35134
3	27	51.9	9	5	ABG46931
4	27	51.9	9	5	ABG60522
5	27	51.9	9	6	ABR56920
6	26	50.0	7	3	AAV84998
7	26	50.0	7	3	ADC44293
8	26	50.0	9	3	AAV63242
9	26	50.0	9	4	AAE07204
10	26	50.0	9	4	AAE24655
11	26	50.0	9	5	ABG35050
12	25	48.1	5	2	AAV95843
13	25	48.1	5	3	AAV84999
14	25	48.1	5	7	ADC44308
15	24	46.2	6	2	AAV03909
16	24	46.2	9	1	AAV10098
17	24	46.2	9	4	ABP21743
18	24	46.2	9	5	AAU94456
19	24	46.2	9	5	AAU94653
20	24	46.2	9	7	ADC44488
21	24	46.2	9	7	ADC59359
22	23	44.2	6	2	AAW85926
23	23	44.2	6	6	ABP99702
24	23	44.2	7	4	AAE64710
25	23	44.2	7	4	AAU04530

26	23	44.2	8	2	AAV58417	AAV58417	Partial p
27	23	44.2	8	2	AAV58416	AAV58416	Partial p
28	23	44.2	8	2	AAV58415	AAV58415	Partial p
29	23	44.2	8	2	AAV58418	AAV58418	Partial p
30	23	44.2	9	2	AAW99327	AAW99327	Human IgG
31	23	44.2	9	2	AAV10322	AAV10322	T cell ep
32	23	44.2	9	2	AAW99853	AAW99853	HIV-1 gpl
33	23	44.2	9	2	AAV47758	AAV47758	Immunogen
34	23	44.2	9	2	AAV45765	AAV45765	Immunogen
35	23	44.2	9	2	AAV45766	AAV45766	Immunogen
36	23	44.2	9	2	AAV48998	AAV48998	Membrane
37	23	44.2	9	3	AAV51358	AAV51358	MHC type
38	23	44.2	9	3	AAV36017	AAV36017	Human bet
39	23	44.2	9	3	AAE27093	AAE27093	Plasmodiu
40	23	44.2	9	4	AAE61826	AAE61826	Human Ig
41	23	44.2	9	5	AAE26894	AAE26894	Linker pe
42	23	44.2	9	5	ABG35060	ABG35060	Angiostat
43	23	44.2	9	5	ABG35128	ABG35128	Pancreat
44	23	44.2	9	5	ABG60516	ABG60516	Selective
45	23	44.2	9	5	ABG80004	ABG80004	MHC class

## ALIGNMENTS

RESULT 1  
AAU04533  
ID AAU04533 standard; peptide; 9 AA.  
XX AC AAU04533;  
XX AC  
DT 26-SEP-2001 (first entry)  
XX  
DE VEGF based monocyclic peptide 11.  
XX  
KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
KW diabetic retinopathy; chronic inflammation; cyclic.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Disulfide-bond 1...9  
FT /note= "This bond cyclises the peptide"  
XX  
XX WO200152875-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 18-JAN-2001; 2001WO-US001533.  
XX  
PR 18-JAN-2000; 2000US-0176293P.  
PR 16-MAY-2000; 2000US-0204590P.  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX Achen MG, Hughes RA, Stacker S, Cendron A;  
PI WPI; 2001-442248/47.  
XX  
PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
from an exposed loop of a growth factor protein by oxidizing the cysteine  
residues.  
PT Claim 49; Page 32; 102pp; English.  
XX  
PS The sequence represents a monomeric monocyclic peptide of the invention,  
whose 3-dimensional structure is modelled on the expose loop of human  
CC VEGF (vascular endothelial growth factor). The invention relates to a  
method of producing a monomeric monocyclic peptide by a measuring beta-  
beta carbon separation distances on opposite antiparallel strands of a

CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CISVPLVPC 9  
 | | | | |  
 Db 1 CISVPLVPC 9

RESULT 2  
 ABG35134  
 ID ABG35134 standard; peptide; 9 AA.

AC ABG35134;

DT 15-JUL-2002 (first entry)

DE Pancreatic islet targeting peptide #10.

XX Targeting peptide; cancer; Hodgkin's disease; cytostatic;  
 KW immunosuppressive; anti-inflammatory; antiarthritic; antiviral;  
 KW antiatherosclerotic; antidiabetic; antibacterial; diabetes mellitus;  
 KW inflammatory disease; arthritis; atherosclerosis; cancer;  
 KW autoimmune disease; bacterial infection; viral infection.

XX Unidentified.

OS WO200220722-A2.

XX 14-MAR-2002.

XX 07-SEP-2001; 2001WO-US027702.

XX 08-SEP-2000; 2000US-0231266P.

XX 17-JAN-2001; 2001US-00765101.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Arap W, Pasqualini R;

XX WPI; 2002-383050/41.

XX Identifying targeting peptides useful for treating e.g. diabetes  
 PT mellitus, inflammatory diseases, cancer, or autoimmune diseases,  
 PT comprises exposing a sample to a phage display library and recovering  
 PT phage bound to the sample.

XX Claim 56; Page 288; 298pp; English.

XX This invention relates to a novel method for identifying disease  
 CC targeting peptides. The method comprises exposing a sample from an organ,  
 CC tissue or cell type of interest, to a phage display library and  
 CC recovering phage bound to the sample (the phage expresses targeting  
 CC peptides). The peptides identified by the method of the invention may  
 CC have cytostatic, immunosuppressive, anti-inflammatory, antiarthritic,  
 CC antiatherosclerotic, antidiabetic, antibacterial and antiviral  
 CC activities. The methods and composition are useful for identifying  
 CC targeting peptides and one or more receptors for a targeting peptide. The  
 CC targeting peptides are used for selective delivery of therapeutic agents,  
 CC including gene therapy vectors and fusion proteins, to specific organs,  
 CC tissues, or cell types in subject. The targeting peptide may also be used  
 CC for treating diseases such as diabetes mellitus, inflammatory diseases,  
 CC arthritis, atherosclerosis, cancer, autoimmune diseases, bacterial and  
 CC viral infections and Hodgkin's disease. The present sequence represents a  
 CC targeting peptide of the invention  
 XX  
 SQ Sequence 9 AA;

Query Match 51.9%; Score 27; DB 5; Length 9;  
 Best Local Similarity 55.8%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CISVPLVPC 9  
 | | | | |  
 Db 1 CMSSPGVAC 9

RESULT 3  
 ABB46931  
 ID ABB46931 standard; peptide; 9 AA.

AC ABB46931;

DT 30-JAN-2002 (first entry)

DE Desmocollin-2 CAR cyclic peptide 13.

XX Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;  
 KW cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;  
 KW organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.

XX Synthetic.

XX WO200172956-A2.

XX 04-OCT-2001.

XX 27-MAR-2001; 2001WO-IB001400.

XX 27-MAR-2000; 2000US-00535852.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Symonds JM, Gour BJ;

XX WPI; 2002-025778/03.

XX Modulating agents for inhibiting or enhancing desmosomal cadherin  
 PT mediated cell adhesion, useful for facilitating wound healing and/or  
 PT reducing scar tissue, treating cancer and inducing apoptosis.

XX Claim 23; Page 111; 127pp; English.

XX The invention relates to modulating agents for inhibiting or enhancing  
 CC desmosomal cadherin mediated cell adhesion, comprising a modulating agent  
 CC comprising a desmosomal cadherin cell adhesion recognition CAR sequence  
 CC (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR  
 CC sequence, a substance such as an antibody or antigen-binding fragment  
 CC that specifically binds a desmosomal cadherin CAR sequence and/or a  
 CC polynucleotide encoding a polypeptide that comprises a desmosomal  
 CC cadherin CAR sequence or analogue. The modulating agents have



CC immunosuppressive, cytostatic and antiapoptotic activity and are used to  
 CC facilitate wound healing and/or reduce scar tissue, for enhancing  
 CC adhesion of foreign tissue implants (e.g. skin graft or organ implant),  
 CC treating an autoimmune blistering disorder and to treat cancer (e.g.  
 CC carcinoma, leukaemia or melanoma) and induce apoptosis  
 XX  
 SQ Sequence 9 AA;

Query Match 51.9%; Score 27; DB 5; Length 9;  
 Best Local Similarity 44.4%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
 Matches 4; Conservative 1; Mismatches 4;

QY 1 CISVPLVPC 9  
 ||:  
 Db 1 CIAFATPC 9

RESULT 4  
 ABG60522  
 ID ABG60522 standard; peptide; 9 AA.  
 XX  
 AC ABG60522;  
 XX

DT 30-JUL-2002 (first entry)

DE Selective targeting peptide #197.

XX Targeting peptide; cancer; arthritis; diabetes; inflammatory disease;  
 KW atherosclerosis; autoimmune disease; bacterial infection; apoptosis;  
 KW viral infection; cardiovascular disease; degenerative disease; ischaemia;  
 KW inflammation; macular degeneration; antiinflammatory; antidiabetic;  
 KW cardiovascular; immunomodulator; antibacterial; antiviral; cytostatic;  
 KW gene therapy.

OS Synthetic.

XX WO200220769-A1.

PN 14-MAR-2002.

XX 07-SEP-2001; 2001WO-US027692.

XX 08-SEP-2000; 2000US-0231266P.

PR 17-JAN-2001; 2001US-00765101.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX Arap W, Pasqualini R;

XX WPI; 2002-415731/44.

XX Targeting peptides identified by phage display, useful for targeting  
 PT delivery to an organ or tissue, particularly for treating a disease, e.g.  
 PT cancer, inflammatory or autoimmune diseases, infections or cardiovascular  
 PT disease.

PS Claim 22; Page 121; 317pp; English.

XX The invention relates to an isolated peptide of 100 amino acids or less  
 CC in size useful for targeting delivery to an organ or tissue, particularly  
 CC for treating a disease, e.g. cancer, arthritis, diabetes, inflammatory  
 CC disease, atherosclerosis, autoimmune disease, bacterial infection, viral  
 CC infection, cardiovascular disease or degenerative disease. The peptide is  
 CC also useful for inducing apoptosis, particularly to a subject with  
 CC ischaemia, cancer, arthritis, diabetes, cardiovascular disease,  
 CC inflammation or macular degeneration. Furthermore, the peptide is useful  
 CC for diagnosing the diseases cited above. Targeting peptides of the  
 CC invention can also be used to deliver an agent to a fetus, by attaching  
 CC a peptide to the agent and administering the peptide to a pregnant  
 CC subject. Sequences ABG60326-ABG60574 represent selective targeting  
 CC peptides of the invention  
 XX

XX Sequence 9 AA;

Query Match 51.9%; Score 27; DB 5; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
 Matches 5; Conservative 1; Mismatches 3;

QY 1 CISVPLVPC 9  
 ||:  
 Db 1 CMSSPGVAC 9

RESULT 5  
 ABR56920  
 ID ABR56920 standard; peptide; 9 AA.  
 XX  
 AC ABR56920;

DT 30-JUL-2003 (first entry)

DE Pancreatic targeting peptide SEQ ID NO:137.

XX Targeting peptide; obesity; lipodystrophy; anorectic; antilipaeamic;  
 KW peptide therapy; gene therapy; infection; human immunodeficiency virus;  
 KW HIV; placental delivery; teratogenic; placenta; adipose; pancreatic;  
 KW beta-3 integrin; beta-5 integrin; spleen; aminopeptidase A.

OS Mus musculus.

OS Synthetic.

XX WO2003022991-A2.

XX 20-MAR-2003.

XX 30-AUG-2002; 2002WO-US027836.

XX 07-SEP-2001; 2001WO-US027692.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX Pasqualini R, Arap W, Kolonin MG;

XX WPI; 2003-371749/35.

XX Treating obesity or a lipodystrophy comprises obtaining a targeting  
 PT peptide selective for adipose tissue, attaching the peptide to a  
 PT therapeutic agent to form a complex, and administering the complex to a  
 PT subject.

XX Example 9; Page 114; 247pp; English.

XX The present invention describes a method for treating obesity or a  
 CC lipodystrophy, which comprises: (a) obtaining a targeting peptide  
 CC selective for adipose tissue; (b) attaching the peptide to a therapeutic  
 CC agent to form a complex; (c) administering the complex to a subject; and  
 CC (d) inducing weight loss in the subject or treating lipodystrophy. The  
 CC adipose targeting peptides have anorectic and antilipaeamic activities,  
 CC and can be used in peptide and gene therapy. The method is used for  
 CC treating obesity or a lipodystrophy that is related to infection with  
 CC human immunodeficiency virus (HIV). The peptides used in the method can  
 CC also be used for targeting delivery to an organ or tissue, such as  
 CC placental delivery. A receptor that binds to a placenta targeting peptide  
 CC is used to screen compounds for teratogenic activity. ABR56806 to  
 CC ABR56927 and ACC79106 to ACC79111 represent sequences used in the  
 CC exemplification of the present invention  
 XX

SQ Sequence 9 AA;

Query Match 51.9%; Score 27; DB 6; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
 Matches 5; Conservative 1; Mismatches 3;

QY 1 CISVPLVPC 9  
 ||:  
 Db 1 CMSSPGVAC 9

```

XX AC ADC44293;
XX DT 18-DEC-2003 (first entry)
XX DE Endothelial cell binding peptide SEQ ID NO:21.
XX KW endothelial cell binding protein; ECBP; anti-tumour; cytostatic;
XX KW vasotropic; antipsoriatic; dermatological; ophthalmological;
XX KW antidiabetic; antiarthritic; vulnerary; antiulcer; antiinflammatory;
XX KW antibacterial; gynaecological; angiogenesis.
XX OS Synthetic.
XX PN WO2003037172-A2.
XX PD 08-MAY-2003.
XX PF 01-NOV-2002; 2002WO-US035258.
XX PR 01-NOV-2001; 2001US-0334822P.
XX PA (GPCB-) GPC BIOTECH INC.
XX PI Gyuris J, Lamphere L, Morris AJ, Tsaioun K;
XX DR WPI; 2003-482072/45.
XX PT Novel synthetic or recombinant polypeptide useful for promoting, reducing
XX PT proliferation and/or migration of endothelial cells, and for modulating
XX PT angiogenesis, has endothelial cell binding protein sequences.
XX PS Claim 3; SEQ ID NO 21; 126pp; English.
XX CC The invention relates to a novel isolated, synthetic or recombinant
XX CC peptide or polypeptide which includes one or more endothelial cell
XX CC binding protein (ECBP) sequences. A peptide of the invention has anti-
XX CC tumour, cytostatic, vasotropic, antipsoriatic, dermatological,
XX CC ophthalmological, antidiabetic, antiarthritic, vulnerary, antiulcer,
XX CC antiinflammatory, antibacterial, and gynaecological activity. The peptide
XX CC is useful for promoting, reducing the proliferation and/or migration of
XX CC endothelial cells, by treating the cells with an ECBP agonist, which is
XX CC preferably the peptide, to promote proliferation and/or migration of the
XX CC treated cells, and for reducing or promoting angiogenesis, by treating
XX CC the cells with an ECBP antagonist, which is preferably the peptide of the
XX CC invention. A peptide of the invention is also useful for manufacturing a
XX CC medicament for promoting angiogenesis, by admixing an ECBP agonist or
XX CC ECBP antagonist to promote or reduce angiogenesis at one or more sites in
XX CC a treated mammal. The medicament is useful for promoting or reducing
XX CC angiogenesis. ECBP sequences are useful to alter the infectivity spectrum
XX CC of a viral particle. The present sequence represents an ECBP of the
XX CC invention.
XX SQ Sequence 7 AA;
    Query Match 50.0%; Score 26; DB 7; Length 7;
    Best Local Similarity 66.7%; Pred. No. 1.4e+06;
    Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLVP 8
DB 2 AIPLVP 7
    :|||||

RESULT 8
AA63242
ID AAY63242 standard; peptide; 9 AA.
XX AC AAY63242;
XX DT 02-MAR-2000 (first entry)
XX DE Protocadherin cell adhesion recognition cyclic peptide SEQ ID NO:2726.

```

---

```

XX AC ADC44293;
XX DT 18-DEC-2003 (first entry)
XX DE Endothelial cell binding peptide SEQ ID NO:21.
XX KW endothelial cell binding protein; ECBP; anti-tumour; cytostatic;
XX KW vasotropic; antipsoriatic; dermatological; ophthalmological;
XX KW antidiabetic; antiarthritic; vulnerary; antiulcer; antiinflammatory;
XX KW antibacterial; gynaecological; angiogenesis.
XX OS Synthetic.
XX PN WO2003037172-A2.
XX PD 08-MAY-2003.
XX PF 01-NOV-2002; 2002WO-US035258.
XX PR 01-NOV-2001; 2001US-0334822P.
XX PA (GPCB-) GPC BIOTECH INC.
XX PI Gyuris J, Lamphere L, Morris AJ, Tsaioun K;
XX DR WPI; 2003-482072/45.
XX PT Novel synthetic or recombinant polypeptide useful for promoting, reducing
XX PT proliferation and/or migration of endothelial cells, and for modulating
XX PT angiogenesis, has endothelial cell binding protein sequences.
XX PS Claim 3; SEQ ID NO 21; 126pp; English.
XX CC The invention relates to a novel isolated, synthetic or recombinant
XX CC peptide or polypeptide which includes one or more endothelial cell
XX CC binding protein (ECBP) sequences. A peptide of the invention has anti-
XX CC tumour, cytostatic, vasotropic, antipsoriatic, dermatological,
XX CC ophthalmological, antidiabetic, antiarthritic, vulnerary, antiulcer,
XX CC antiinflammatory, antibacterial, and gynaecological activity. The peptide
XX CC is useful for promoting, reducing the proliferation and/or migration of
XX CC endothelial cells, by treating the cells with an ECBP agonist, which is
XX CC preferably the peptide, to promote proliferation and/or migration of the
XX CC treated cells, and for reducing or promoting angiogenesis, by treating
XX CC the cells with an ECBP antagonist, which is preferably the peptide of the
XX CC invention. A peptide of the invention is also useful for manufacturing a
XX CC medicament for promoting angiogenesis, by admixing an ECBP agonist or
XX CC ECBP antagonist to promote or reduce angiogenesis at one or more sites in
XX CC a treated mammal. The medicament is useful for promoting or reducing
XX CC angiogenesis. ECBP sequences are useful to alter the infectivity spectrum
XX CC of a viral particle. The present sequence represents an ECBP of the
XX CC invention.
XX SQ Sequence 7 AA;
    Query Match 50.0%; Score 26; DB 3; Length 7;
    Best Local Similarity 66.7%; Pred. No. 1.4e+06;
    Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLVP 8
DB 2 AIPLVP 7
    :|||||

RESULT 7
ADC44293
ID ADC44293 standard; peptide; 7 AA.

```

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; F-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1. .9  
 XX  
 XX WO957149-A2.  
 XX  
 XX 11-NOV-1999.  
 XX  
 XX 05-MAY-1999; 99WO-CA000363.  
 XX  
 PR 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 XX Blaschuk OM, Gour BU, Byers S;  
 PI WPI; 2000-038791/03.  
 DR  
 XX  
 XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT -mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX  
 XX Claim 84; Page 201; 252pp; English.  
 XX  
 XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MA's can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 XX Sequence 9 AA;  
 CC  
 CC Query Match 50.0%; Score 26; DB 3; Length 9;  
 CC Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
 CC Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 C1SVPLVPC 9  
 DB | : : | |  
 1 CFAOLDLVC 9

RESULT 9  
 AAE07204  
 ID AAE07204 standard; peptide; 9 AA.  
 XX  
 AC AAE07204;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Modified colostrinin cyclic peptide #10.  
 XX  
 KW Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;  
 KW Parkinson's disease; Alzheimer's disease; mental disorder; food additive;  
 KW central nervous system disorder; neurodegenerative disorder; weight loss;  
 KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;  
 KW acquired immunological deficiency; neurological disorder; dementia;  
 KW antiviral; cyclic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl; this residue forms a cyclic  
 XX linkage with Ser found at the C-terminal end"  
 PN WO20015199-A1.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 26-JAN-2001; 2001WO-GB000329.  
 XX  
 XX 26-JAN-2000; 2000GB-00001825.  
 XX  
 XX (REG-) REGEN THERAPEUTICS PLC.  
 XX  
 XX Georgiades JA;  
 XX  
 XX WPI; 2001-488775/53.  
 XX  
 XX Peptide useful as an interval in the treatment of e.g. disorders of the  
 PT immune system and the central nervous system comprises ten amino-terminal  
 PT amino acid sequence derived from peptides present in colostrinin.  
 XX  
 XX Example 2; Page 9; 40pp; English.  
 XX  
 CC The invention relates to colostrinin peptide fragments which are useful,  
 CC inter alia, in the treatment of chronic disorders of the immune system  
 CC and the central nervous system. Colostrinin peptides are used as a  
 CC medicament in the treatment of neurological disorders e.g., dementia,  
 CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron  
 CC disease e.g., Parkinson's disease, mental disorders e.g. psychosis and  
 CC neurosis, in acquired immunological deficiencies, chronic bacterial and  
 CC viral infections and diseases characterised by the presence of beta-  
 CC amyloid plaques and as a dietary supplement for babies, small children,  
 CC adults and senile persons, who have been subjected to chemotherapy or  
 CC have suffered from cachexia or weight loss due to the chronic disease.  
 CC Colostrinin peptides are also used as food additives and as an auxiliary  
 CC withdrawal treatment for drug addicts, after a period of detoxification  
 CC and in persons dependent on stimulants. Colostrinin peptides are used to  
 CC prepare antibodies and to treat emotional disturbances, e.g. emotional  
 CC disturbances of psychiatric patients in a state of depression. These  
 CC colostrinin peptides improves the development of immune system in a new  
 CC born child and to correct the immunological deficiencies in a child. The  
 CC present sequence is modified colostrinin cyclic peptide #10 related to  
 CC the invention  
 XX  
 XX Sequence 9 AA;  
 CC  
 CC Query Match 50.0%; Score 26; DB 4; Length 9;  
 CC Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 CC Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 C1SVPLV 7

Db 1 CLPLPLV 7  
 : : |||

RESULT 10  
 AAM24655  
 ID AAM24655 standard; peptide; 9 AA.

XX AC AAM24655;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE Human MHC class I molecule HLA-A2 binding 83P5G4 peptide #32.  
 XX KW 83P5G4-related protein; prostate; testis; tissue; cancer; bladder; liver;  
 XX KW tumour; kidney; brain; bone; ovary; breast; pancreas; colon; lung; serum;  
 XX KW cytostatic; gene therapy; antibody therapy; ribozyme; blood; cervix;  
 XX KW single chain monoclonal antibody; urine; uterus; rectum; stomach; human;  
 XX KW chromosome 1q31-q32.

XX OS Homo sapiens.

XX PN W0200159115-A2.

XX PD 16-AUG-2001.

XX PF 09-FEB-2001; 2001WO-US004426.

XX PR 09-FEB-2000; 2000US-0191261P.

XX PA (UROG-) UROGENESYS INC.

XX PI Hubert RS, Afar DEH, Challita-Eid PM, Paris M, Levin E;

XX PI Mitchell SC, Jakobovits A;

XX DR WPI; 2001-514669/56.

XX PT An isolated 83P5G4-related protein useful as a diagnostic and/or  
 PT therapeutic agent in multiple cancers such as prostate, bladder and bone  
 PT cancer.

XX PS Example 15; Page 77; 112pp; English.

XX CC The polypeptide sequences represent the 83P5G4-related protein and  
 CC peptide fragments of the protein. 83P5G4 exhibits prostate specific  
 CC expression in normal adult tissue, but it is also aberrantly expressed in  
 CC many cancers including tumours of the prostate, testis, bladder, kidney,  
 CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,  
 CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and  
 CC peptide fragments and specific PCR primers are therefore useful for  
 CC diagnosing and treating cancer. A vector comprising a polynucleotide  
 CC which encodes a single chain monoclonal antibody, that immunospecifically  
 CC binds to an 83P5G4-related protein, and a ribozyme capable of cleaving a  
 CC polynucleotide having the 83P5G4 coding sequence, are both useful in the  
 CC preparation of a composition for treating a patient with a cancer that  
 CC expresses 83P5G4. The sequences can be used in diagnostic methods to  
 CC monitor the level of 83P5G4 gene products in serum, blood, urine and  
 CC tissue and to thereby detect the presence of cancerous cells

XX SQ Sequence 9 AA;

Query Match 50.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ISVPLVPC 9  
 : : |||

Db 2 LPLPLRPC 9

RESULT 11  
 ABG35050  
 ID ABG35050 standard; peptide; 9 AA.

XX AC ABG35050;  
 XX DT 15-JUL-2002 (first entry)  
 XX DE Endostatin targeting peptide #49.  
 XX KW Targeting peptide; cancer; Hodgkin's disease; cytostatic;  
 XX KW immunosuppressive; anti-inflammatory; antiarthritic; antiviral;  
 XX KW antiatherosclerotic; antidiabetic; antibacterial; diabetes mellitus;  
 XX KW inflammatory disease; arthritis; atherosclerosis; cancer;  
 XX KW autoimmune disease; bacterial infection; viral infection.  
 XX OS Unidentified.  
 XX PN W0200220722-A2.  
 XX PD 14-MAR-2002.  
 XX PF 07-SEP-2001; 2001WO-US027702.  
 XX PR 08-SEP-2000; 2000US-0231266P.  
 XX PR 17-JAN-2001; 2001US-00765101.  
 XX PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX PI Arap W, Pasqualini R;  
 XX DR WPI; 2002-383050/41.  
 XX PT Identifying targeting peptides useful for treating e.g. diabetes  
 PT mellitus, inflammatory diseases, cancer, or autoimmune diseases,  
 PT comprises exposing a sample to a phage display library and recovering  
 PT phage bound to the sample.

XX Claim 56; Page 252; 298pp; English.

XX CC This invention relates to a novel method for identifying disease  
 CC targeting peptides. The method comprises exposing a sample from an organ,  
 CC tissue or cell type of interest, to a phage display library and  
 CC recovering phage bound to the sample (the phage expresses targeting  
 CC peptides). The peptides identified by the method of the invention may  
 CC have cytostatic, immunosuppressive, anti-inflammatory, antarthritic,  
 CC antiatherosclerotic, antidiabetic, antibacterial and antiviral  
 CC activities. The methods and composition are useful for identifying  
 CC targeting peptides and one or more receptors for a targeting peptide. The  
 CC targeting peptides are used for selective delivery of therapeutic agents,  
 CC including gene therapy vectors and fusion proteins, to specific organs,  
 CC tissues, or cell types in subject. The targeting peptide may also be used  
 CC for treating diseases such as diabetes mellitus, inflammatory diseases,  
 CC arthritis, atherosclerosis, cancer, autoimmune diseases, bacterial and  
 CC viral infections and Hodgkin's disease. The present sequence represents a  
 CC targeting peptide of the invention

XX SQ Sequence 9 AA;

Query Match 50.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLVPC 9  
 : : |||

Db 1 CRSIPPVRC 9

RESULT 12  
 AAR35843  
 ID AAR35843 standard; peptide; 5 AA.

XX AC AAR35843;

XX DT 26-JUN-1996 (first entry)

XX

DE IGF-1 mutein C67 residues 63-67.  
 XX Insulin-like growth factor-1; IGF-1; polyethylene glycol; PEG; triflate;  
 KW IGF-1/PEG conjugate; maleimide; sulphhydryl; thiol; tresylate; aziride;  
 KW exirane; 5-pyridyl; therapy; dwarfism; diabetes; periodontal disease;  
 KW osteoporosis.  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 5 /note= "A5C"  
 FT  
 XX WO9532003-A1.  
 XX 30-NOV-1995.  
 XX  
 XX 24-MAY-1995; 95WO-US006540.  
 XX 24-MAY-1994; 94US-00248273.  
 XX (AMGE-) AMGEN BOULDER INC.  
 XX Cox GN, Medermott MJ, Ko C;  
 XX WPI; 1996-020360/02.  
 XX Conjugates for treatment of, e.g. dwarfism, diabetes, or osteoporosis -  
 PT comprising polyethylene glycol attached to mutein of IGF at free  
 PT cysteine.  
 XX Example 2; Page 22; 48pp; English.  
 XX AAR95832-R95844 represent the altered fragments of the insulin-like  
 CC growth factor-1 (IGF-1) muteins of the invention. This sequence  
 CC represents the N-terminal residues 63-67 of the IGF-1 mutein C67. The  
 CC wild type IGF-1 sequence is represented by AAR87744. These muteins  
 CC contain a non-native cysteine residue substituted for one of the first (or  
 CC last) four amino acid residues of this sequence. Polyethylene glycol  
 CC (PEG) conjugates are then created from the muteins, where the PEG is  
 CC attached to the non-native cysteine residue. The PEG is attached to the  
 CC free cysteine through an activating group selected from maleimide,  
 CC sulphhydryl, thiol, triflate, tresylate, aziride, exirane or 5-pyridyl.  
 CC The conjugates can also comprise a second polypeptide attached to the  
 CC PEG. The conjugates may be used for the treatment of IGF associated  
 CC conditions, such as dwarfism, diabetes, periodontal disease or  
 CC osteoporosis. Advantages associated with these conjugates are that they  
 CC have a higher molecular weight, and an extended circulating half life in  
 CC comparison to wild type IGF  
 XX Sequence 5 AA;  
 SQ  
 Query Match 48.1%; Score 25; DB 2; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 PLVPC 9  
 Db |||||  
 1 PLKPC 5  
 RESULT 13  
 ID AAY84999 standard; peptide; 5 AA.  
 XX  
 AC AAY84999;  
 XX  
 DT 21-AUG-2000 (first entry)  
 DE Amino acid sequence of an endothelial cell-binding peptide motif.  
 XX Endothelial cell-binding peptide; Genetic display package;  
 KW peptide display library; affinity selection; population display package;  
 KW

KW cell proliferation; cell differentiation; cell death; cell migration;  
 KW angiogenic activity; infective peptide; anti-fungal; anti-bacterial;  
 KW receptor protein effector.  
 XX Unidentified.  
 OS  
 XX WO200023465-A2.  
 XX  
 XX 27-APR-2000.  
 PD  
 XX  
 XX 19-OCT-1999; 99WO-US024276.  
 PF  
 XX 19-OCT-1998; 98US-00174943.  
 PR  
 XX (MITO-) MITOTIX INC.  
 PA  
 XX Gyuris J, Morris AJ;  
 PI WPI; 2000-339649/29.  
 XX  
 DR  
 XX  
 XX Generating a peptide with a selected biological activity useful for  
 PT identifying endothelial inhibitors and peptides with anti-angiogenic  
 PT activity by combining peptide display libraries in a display and a  
 PT secretion mode.  
 PT  
 XX Disclosure; Page 42; 86pp; English.  
 PS  
 XX The present sequence represents an endothelial cell-binding peptide  
 CC motif. The peptide may be identified using the method of the invention.  
 CC The specification describes a method for generating a peptide having a  
 CC selected biological activity. The method comprises displaying the  
 CC peptides on the outer surface of a genetic display package to create a  
 CC peptide display library, and using affinity selection to enrich the  
 CC population display packages for those containing peptides which have  
 CC desired specificity to the target cell. The method may be used in the  
 CC selection of peptides having effects on cell proliferation,  
 CC differentiation, death and migration, as well as in the identification of  
 CC peptides which have anti-proliferative activity with respect to one or  
 CC more types of cells, peptides with (anti-)angiogenic activity, anti-  
 CC infective peptides (e.g. which are active as anti-fungal or anti-  
 CC bacterial), receptor protein effectors, and ligands for orphan receptors  
 CC for which no ligand is known. Moreover, the method may be used to test  
 CC functional ligand-receptor or ligand-ion channel interactions for cell  
 CC surface-localized receptors and channels  
 XX Sequence 5 AA;  
 SQ  
 Query Match 48.1%; Score 25; DB 3; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 VPLVP 8  
 Db :|||  
 1 IPLVP 5  
 RESULT 14  
 ID ADC44308 standard; peptide; 5 AA.  
 XX  
 AC ADC44308;  
 XX  
 DT 18-DEC-2003 (first entry)  
 DE Endothelial cell binding peptide SEQ ID NO:36.  
 XX Endothelial cell binding protein; EGBP; anti-tumour; cytostatic;  
 KW vasotropic; antipsoriatic; dermatological; ophthalmological;  
 KW antidiabetic; antiarthritic; vulnary; antituber; antiinflammatory;  
 KW antibacterial; gynaecological; angiogenesis.  
 XX Synthetic.  
 OS

PN WO2003037172-A2.  
 XX  
 PD  
 XX  
 XX 08-MAY-2003.  
 XX  
 PF 01-NOV-2002; 2002WO-US035258.  
 XX  
 XX  
 PR 01-NOV-2001; 2001US-0334822P.  
 XX  
 XX  
 PA (GPCB-) GPC BIOTECH INC.  
 XX  
 PI Gyuris J, Lamphere L, Morris AJ, Tsaoun K;  
 XX  
 XX WPI; 2003-482072/45.  
 DR  
 XX  
 XX Novel synthetic or recombinant polypeptide useful for promoting, reducing  
 PT proliferation and/or migration of endothelial cells, and for modulating  
 PT angiogenesis, has endothelial cell binding protein sequences.  
 XX  
 XX Claim 3; SEQ ID NO 36; 126pp; English.  
 PS  
 XX The invention relates to a novel isolated, synthetic or recombinant  
 CC peptide or polypeptide which includes one or more endothelial cell  
 CC binding protein (ECBP) sequences. A peptide of the invention has anti-  
 CC tumour, cytostatic, vasotropic, antipsoriatic, dermatological,  
 CC ophthalmological, antidiabetic, antiarthritic, vulnary, antitumor,  
 CC antiinflammatory, antibacterial, and gynaecological activity. The peptide  
 CC is useful for promoting, reducing the proliferation and/or migration of  
 CC endothelial cells, by treating the cells with an ECBP agonist, which is  
 CC preferably the peptide, to promote proliferation and/or migration of the  
 CC treated cells, and for reducing or promoting angiogenesis, by treating  
 CC the cells with an ECBP antagonist, which is preferably the peptide of the  
 CC invention. A peptide of the invention is also useful for manufacturing a  
 CC medicament for promoting angiogenesis, by admixing an ECBP agonist or  
 CC ECBP antagonist to promote or reduce angiogenesis at one or more sites in  
 CC a treated mammal. The medicament is useful for promoting or reducing  
 CC angiogenesis. ECBP sequences are useful to alter the infectivity spectrum  
 CC of a viral particle. The present sequence represents an ECBP of the  
 CC invention.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 49.1%; Score 25; DB 7; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 VPLVP 8  
 Db :||||  
 1 IPLVP 5  
 RESULT 15  
 AAY03909  
 ID AAY03909 standard; peptide; 6 AA.  
 XX  
 AC AAY03909;  
 XX  
 DT 08-JUN-1999 (first entry)  
 XX  
 DE Hepatitis C inhibitor peptide.  
 XX  
 KW Hepatitis C; inhibitor; NS3 protease; NS4A cofactor.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-Ac-Asp"  
 XX  
 PN WO9907733-A2.  
 XX  
 PD 18-FEB-1999.  
 XX  
 XX 10-AUG-1998; 98WO-CA000765.

XX 11-AUG-1997; 97US-0055186P.  
 PR  
 XX (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.  
 PA  
 XX Llinas-Brunet M, Poupart M, Rancourt J, Simoneau B, Tsantrizos Y;  
 PI Wernic D;  
 PR  
 XX WPI; 1999-167361/14.  
 DR  
 XX New peptides inhibitors of NS3 protease - useful for treatment of  
 PT hepatitis C infections.  
 PT  
 XX Disclosure; Page 108; 158pp; English.  
 PS  
 XX New generically described peptides are disclosed which are NS3 protease  
 CC inhibitors useful for treating hepatitis C virus (HCV) infection. The  
 CC present sequence represents a specific example of these peptides. In a  
 CC test to determine inhibition in a NS3 protease/NS4A cofactor peptide  
 CC radiometric assay, this peptide had an IC50 value of 92 micromolar  
 CC  
 XX Sequence 6 AA;  
 SQ  
 Query Match 46.2%; Score 24; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LVPC 9  
 Db :||||  
 3 LVPC 6  
 Search completed: September 5, 2004, 11:03:36  
 Job time : 39.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:05:54 ; Search time 35,6667 Seconds  
(without alignments)  
79,502 Million cell updates/sec

Title: US-09-761-636A-14  
Perfect score: 52  
Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 119143

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	9	9	US-09-761-636A-14
2	27	51.9	9	12	US-10-363-208-242
3	27	51.9	9	15	US-10-154-884B-11221
4	27	51.9	9	15	US-10-154-884B-11228
5	27	51.9	9	15	US-10-154-884B-11250
6	27	51.9	9	15	US-10-154-884B-11254
7	27	51.9	9	15	US-10-154-884B-11258
8	27	51.9	9	15	US-10-154-884B-11259
9	27	51.9	9	15	US-10-154-884B-11265
10	27	51.9	9	15	US-10-154-884B-11275
11	27	51.9	9	15	US-10-154-884B-11280
12	26	50.0	7	14	US-10-286-457-21
13	26	50.0	9	9	US-09-780-053-148
14	26	50.0	9	12	US-10-363-208-137
15	26	50.0	9	14	US-10-006-869-2726

16	26	50.0	9	15	US-10-395-032-2726
17	25	48.1	5	14	US-10-286-457-36
18	24.5	47.1	9	15	US-10-428-335-52
19	24	46.2	9	10	US-09-932-165-439
20	24	46.2	9	10	US-09-932-165-636
21	24	46.2	9	14	US-10-286-457-216
22	24	46.2	9	15	US-10-428-335-90
23	23	44.2	7	9	US-09-761-636A-11
24	23	44.2	9	9	US-09-997-579-19
25	23	44.2	9	12	US-10-363-208-147
26	23	44.2	9	12	US-10-363-208-236
27	23	44.2	9	14	US-10-038-407-29
28	23	44.2	9	14	US-10-254-466A-123
29	23	44.2	9	14	US-10-286-457-165
30	23	44.2	9	14	US-10-286-457-500
31	23	44.2	9	14	US-10-164-279-15
32	23	44.2	9	16	US-10-777-053-304
33	22	42.3	4	9	US-09-947-387-69
34	22	42.3	4	15	US-10-138-375-69
35	22	42.3	5	9	US-09-947-387-68
36	22	42.3	5	9	US-09-947-387-116
37	22	42.3	5	15	US-10-138-375-68
38	22	42.3	5	15	US-10-138-375-116
39	22	42.3	6	9	US-09-947-387-67
40	22	42.3	6	9	US-09-947-387-115
41	22	42.3	6	14	US-10-010-184A-2
42	22	42.3	6	15	US-10-138-375-67
43	22	42.3	6	15	US-10-138-375-115
44	22	42.3	7	9	US-09-947-387-114
45	22	42.3	7	12	US-10-609-217-302

ALIGNMENTS

RESULT 1  
US-09-761-636A-14  
; Sequence 14, Application US/09761636A  
; Patent No. US20020065218A1  
; GENERAL INFORMATION:  
; APPLICANT: ACHEN, Marc  
; APPLICANT: STACKER, Steven  
; APPLICANT: HUGHES, Richard  
; APPLICANT: CENDRON, Angela  
; TITLE OF INVENTION: VEGF-D/VEGF-C/VEGF PEPTIDOMIMETIC INHIBITOR  
; FILE REFERENCE: 1064/48505 Achen et al  
; CURRENT APPLICATION NUMBER: US/09/761,636A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/176,293  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/204,590  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 14  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-636A-14

Query Match 100.0%; Score 52; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CISVPLVPC 9  
Db 1 CISVPLVPC 9

RESULT 2  
US-10-363-208-242  
; Sequence 242, Application US/10363208  
; Publication No. US20040048243A1

GENERAL INFORMATION:  
APPLICANT: Board of Regents, The University of Texas System  
TITLE OF INVENTION: Methods and Compositions for In Vitro Targeting

FILE REFERENCE: 005774.P005PCT

CURRENT APPLICATION NUMBER: US/10/363,208

CURRENT FILING DATE: 2003-03-07

NUMBER OF SEQ ID NOS: 273

SOFTWARE: PatentIn version 3.1

SEQ ID NO 242

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: Peptide

LOCATION: (1)...(9)

OTHER INFORMATION: synthetic construct

US-10-363-208-242

Query Match 51.9%; Score 27; DB 12; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.2e+06;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 C1SVPLVPC 9

|:|:|:|

Db 1 CMSSPGVAC 9

#### RESULT 3

US-10-154-884B-11221

Sequence 11221, Application US/10154884B

Publication No. US20040005561A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc W.

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

TITLE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-013521US

CURRENT APPLICATION NUMBER: US/10/154,884B

CURRENT FILING DATE: 2002-05-23

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/222,903

PRIOR FILING DATE: 2000-08-03

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 11290

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 11221

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapiens

US-10-154-884B-11221

Query Match 51.9%; Score 27; DB 15; Length 9;  
Best Local Similarity 56.7%; Pred. No. 1.2e+06;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 C1SVPL 6

|:|:|:|

Db 2 CLSVPV 7

#### RESULT 4

US-10-154-884B-11228

Sequence 11228, Application US/10154884B

Publication No. US20040005561A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc W.

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

TITLE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-013521US

CURRENT APPLICATION NUMBER: US/10/154,884B

CURRENT FILING DATE: 2002-05-23

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/222,903

PRIOR FILING DATE: 2000-08-03

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 11290

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 11228

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapiens

US-10-154-884B-11228

Query Match 51.9%; Score 27; DB 15; Length 9;

Best Local Similarity 66.7%; Pred. No. 1.2e+06;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 C1SVPL 6

|:|:|:|

Db 4 CLSVPV 9

#### RESULT 5

US-10-154-884B-11250

Sequence 11250, Application US/10154884B

Publication No. US20040005561A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc W.

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

TITLE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-013521US



```
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11250
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11250

Query Match          51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CUSVPL 6
   |:|:|:
Db 4 CLSVFPV 9

RESULT 6
US-10-154-884B-11254
; Sequence 11254, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11254
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11258

Query Match          51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CUSVPL 6
   |:|:|:
Db 4 CLSVFPV 9

RESULT 7
US-10-154-884B-11258
; Sequence 11258, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11258
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11258

Query Match          51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CUSVPL 6
   |:|:|:
Db 4 CLSVFPV 9
```

```
; CURRENT APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11254
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11254

Query Match          51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CUSVPL 6
   |:|:|:
Db 4 CLSVFPV 9

RESULT 7
US-10-154-884B-11258
; Sequence 11258, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11258
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11258

Query Match          51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CUSVPL 6
   |:|:|:
Db 4 CLSVFPV 9
```

## RESULT 8

US-10-154-884B-11259  
; Sequence 11259, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US/10/154,884B  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 11259  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-154-884B-11259

Query Match 51.9%; Score 27; DB 15; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 C1SVPL 6  
|:|:|:  
Db 4 CLSVPV 9

## RESULT 9

US-10-154-884B-11265  
; Sequence 11265, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US/10/154,884B  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545

; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 11265  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-154-884B-11265

Query Match 51.9%; Score 27; DB 15; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 C1SVPL 6  
|:|:|:  
Db 4 CLSVPV 9

## RESULT 10

US-10-154-884B-11275  
; Sequence 11275, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US/10/154,884B  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 11275  
; LENGTH: 9

```
;
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11275

Query Match      51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTSVPL 6
DB 4 CLSVFP 9

RESULT 11
US-10-154-884B-11280
; Sequence 11280, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11280
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-884B-11280

Query Match      51.9%; Score 27; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTSVPL 6
DB 4 CLSVFP 9

RESULT 12
US-10-286-457-21
; Sequence 21, Application US/10286457
; Publication No. US20030166004A1
; GENERAL INFORMATION:
; APPLICANT: JENO GYURIS et al.
; TITLE OF INVENTION: ENDOTHELIAL-CELL BINDING PEPTIDES FOR DIAGNOSIS AND THERAPY
```

```
;
; FILE REFERENCE: GPCI-P01-178
; CURRENT APPLICATION NUMBER: US/10/286,457
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/334822
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 684
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence isolated from random peptide libraries, based
; OTHER INFORMATION: ability to selectively bind to endothelial cells
US-10-286-457-21

Query Match      50.0%; Score 26; DB 14; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLVP 8
DB 2 AIPLVP 7

RESULT 13
US-09-780-053-148
; Sequence 148, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-148

Query Match      50.0%; Score 26; DB 9; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISVPLVPC 9
DB 2 LPLPLRPC 9

RESULT 14
US-10-363-208-137
; Sequence 137, Application US/10363208
; Publication No. US20040048243A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: Methods and Compositions for In Vitro Targeting
; FILE REFERENCE: 005774.P005PCT
; CURRENT APPLICATION NUMBER: US/10/363,208
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 137
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(9)
; OTHER INFORMATION: synthetic construct
US-10-363-208-137

Query Match      50.0%; Score 26; DB 12; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.2e+06;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTSVPLVPC 9
Db      1 CRSLPPVRC 9

RESULT 15
US-10-006-869-2726
; Sequence 2726, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2726
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: Protocadherin cell adhesion recognition sequence
US-10-006-869-2726

Query Match      50.0%; Score 26; DB 14; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.2e+06;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTSVPLVPC 9
Db      1 CPALDLVTC 9
```

Search completed: September 5, 2004, 11:16:00  
Job time : 35.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 11:01:38 ; Search time 12 Seconds  
(without alignments)  
38.719 Million cell updates/sec

Title: US-09-761-636A-14  
Perfect score: 52  
Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 89883

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	51.9	9	4	US-09-535-852-1681
2	26	50.0	9	4	US-09-187-859-2726
3	26	50.0	9	4	US-09-839-542B-2726
4	24	46.2	7	1	US-08-261-206A-18
5	23	44.2	8	1	US-08-189-331-147
6	23	44.2	8	1	US-08-189-331-148
7	23	44.2	8	1	US-08-189-331-149
8	23	44.2	8	1	US-08-189-331-150
9	23	44.2	8	1	US-08-633-760-9
10	23	44.2	8	2	US-08-471-068-147
11	23	44.2	8	2	US-08-471-068-148
12	23	44.2	8	2	US-08-471-068-149
13	23	44.2	8	2	US-08-471-068-150
14	23	44.2	9	3	US-08-891-271-4
15	23	44.2	9	3	US-09-258-754-373
16	23	44.2	9	3	US-09-042-107-373
17	23	44.2	9	4	US-09-997-579-19
18	23	44.2	9	4	US-09-722-250D-373
19	22	42.3	4	3	US-09-357-952-69
20	22	42.3	4	3	US-09-521-650-69
21	22	42.3	4	4	US-09-168-888-69
22	22	42.3	5	3	US-09-357-952-68
23	22	42.3	5	3	US-09-357-952-116
24	22	42.3	5	4	US-09-521-650-68
25	22	42.3	5	4	US-09-521-650-116
26	22	42.3	5	4	US-09-168-888-68
27	22	42.3	5	4	US-09-168-888-116

```

28 22 42.3 6 3 US-09-357-952-67 Sequence 67, Appl
29 22 42.3 6 3 US-09-357-952-115 Sequence 115, App
30 22 42.3 6 4 US-09-521-650-67 Sequence 67, Appl
31 22 42.3 6 4 US-09-521-650-115 Sequence 115, App
32 22 42.3 6 4 US-09-168-888-67 Sequence 67, Appl
33 22 42.3 6 4 US-09-168-888-115 Sequence 115, App
34 22 42.3 7 1 US-08-261-206A-22 Sequence 22, Appl
35 22 42.3 7 3 US-09-357-952-114 Sequence 114, App
36 22 42.3 7 4 US-09-521-650-114 Sequence 114, App
37 22 42.3 7 4 US-09-168-888-114 Sequence 114, App
38 22 42.3 7 4 US-09-428-082B-302 Sequence 302, App
39 22 42.3 8 3 US-09-357-952-113 Sequence 113, App
40 22 42.3 8 4 US-09-521-650-113 Sequence 113, App
41 22 42.3 8 4 US-09-168-888-113 Sequence 113, App
42 22 42.3 9 2 US-08-456-670B-26 Sequence 26, Appl
43 22 42.3 9 3 US-09-258-754-249 Sequence 249, App
44 22 42.3 9 3 US-08-660-092-123 Sequence 123, App
45 22 42.3 9 3 US-08-660-092-126 Sequence 126, App

```

## ALIGNMENTS

```

RESULT 1
US-09-535-852-1681
; Sequence 1681, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1681
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cyclicized modulating agent comprising
; OTHER INFORMATION: desmocollin-2 cell adhesion recognition sequence
US-09-535-852-1681

```

```

Query Match 51.9%; Score 27; DB 4; Length 9;
Best Local Similarity 44.4%; Pred No. 3e+05;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CISVPLVPC 9
|||
Db 1 CIAFATTPC 9

```

```

RESULT 2
US-09-187-859-2726
; Sequence 2726, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2726
; LENGTH: 9
; TYPE: PRT

```



Query Match 44.2%; Score 23; DB 1; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISVP 5  
|:|  
Db 4 CVSAP 8

## RESULT 6

US-08-189-331-148  
; Sequence 148, Application US/08189331  
; Patent No. 5747334  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/189,331  
; FILING DATE: Concurrently herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 148:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-189-331-148

Query Match 44.2%; Score 23; DB 1; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISVP 5  
|:|  
Db 3 CVSAP 7

## RESULT 7

US-08-189-331-149  
; Sequence 149, Application US/08189331  
; Patent No. 5747334  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas

; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/189,331  
; FILING DATE: Concurrently herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 149:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-189-331-149

Query Match 44.2%; Score 23; DB 1; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISVP 5  
|:|  
Db 2 CVSAP 6

## RESULT 8

US-08-189-331-150  
; Sequence 150, Application US/08189331  
; Patent No. 5747334  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/189,331  
; FILING DATE: Concurrently herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 150:

SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-189-331-150

Query Match 44.2%; Score 23; DB 1; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVLP 5  
|:|:  
Db 1 CVSAP 5

## RESULT 9

US-08-633-760-9  
; Sequence 9, Application US/08633760  
; Patent No. 5804429  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: FUJIMURA, TAKAO  
; APPLICANT: ISHII, YOSHINORI  
; APPLICANT: NOGUCHI, YUJI  
; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/633,760  
FILING DATE: 01-MAY-1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-929-0 PCT  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-633-760-9

Query Match 44.2%; Score 23; DB 1; Length 8;  
Best Local Similarity 42.9%; Pred. No. 3e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVPLV 7  
|:|:  
Db 2 CAAVPML 8

## RESULT 10

US-08-471-068-147  
; Sequence 147, Application US/08471068  
; Patent No. 5948635  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowlkes, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,068  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/189,331  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,972  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 147:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-471-068-147

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVLP 5  
|:|:  
Db 4 CVSAP 8

## RESULT 11

US-08-471-068-148  
; Sequence 148, Application US/08471068  
; Patent No. 5948635  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowlkes, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,068  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/189,331  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-155  
TELEPHONE: 212 790-9090  
TELEFAX: 212 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-471-068-148

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CUSVP 5  
Db 3 CVSAP 7

RESULT 12  
US-08-471-068-149  
Sequence 149, Application US/08471068  
Patent No. 5948635  
GENERAL INFORMATION:  
APPLICANT: Kay, B. K.  
TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
NUMBER OF SEQUENCES: 186  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,068  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/189,331  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-155  
TELEPHONE: 212 790-9090  
TELEFAX: 212 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 149:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-471-068-149

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CUSVP 5  
Db 2 CVSAP 6

RESULT 13  
US-08-471-068-150  
Sequence 150, Application US/08471068  
Patent No. 5948635  
GENERAL INFORMATION:  
APPLICANT: Kay, B. K.  
TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
NUMBER OF SEQUENCES: 186  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,068  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/189,331  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-155  
TELEPHONE: 212 790-9090  
TELEFAX: 212 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 150:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-471-068-150

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CUSVP 5  
Db 1 CVSAP 5

RESULT 14  
US-08-891-271-4  
Sequence 4, Application US/08891271  
Patent No. 6165476

GENERAL INFORMATION:  
APPLICANT: Strom, Terry B.  
APPLICANT: Syckowski, Arthur J.  
APPLICANT: Zheng, Xin Xiao  
TITLE OF INVENTION: Fusion Proteins with an Immunoglobulin  
TITLE OF INVENTION: Hing Region Linker  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/891,271  
FILING DATE: 10-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hogle, Doreen M.  
REGISTRATION NUMBER: 36,361  
REFERENCE/DOCKET NUMBER: BIH97-07  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-891-271-4

Query Match 44.2%; Score 23; DB 3; Length 9;  
Best Local Similarity 44.4%; Pred. No. 3e+05; 2; Indels 2; Gaps 1;  
Matches 4; Conservative 1; Mismatches 2;

QY 1 C1SVPLVPC 9  
Db 2 CVECP--PC 8

RESULT 15  
US-09-258-754-373  
Sequence 373, Application US/09258754  
Patent No. 6174687  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
APPLICANT: Pasqualini, Renata  
APPLICANT: Rajotte, Daniel  
TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
TITLE OF INVENTION: Membrane Dipeptidase  
FILE REFERENCE: P-LJ 3443  
CURRENT APPLICATION NUMBER: US/09/258,754  
CURRENT FILING DATE: 1999-02-26  
EARLIER APPLICATION NUMBER: 09/042,107  
EARLIER FILING DATE: 1998-03-13  
NUMBER OF SEQ ID NOS: 452  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 373  
LENGTH: 9  
TYPE: PPT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-258-754-373

Query Match 44.2%; Score 23; DB 3; Length 9;  
Best Local Similarity 80.0%; Pred. No. 3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 PLVPC 9  
Db 4 PLVAC 8  
Search completed: September 5, 2004, 11:07:05  
Job time : 12 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 09:47:29 ; Search time 9,9798 Seconds  
(without alignments)  
125.302 Million cell updates/sec

Title: US-09-761-636A-5

Perfect score: 72

Sequence: 1 CASELGKSTNTFC 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	59.7	104	A39035	ribonuclease-relat
2	41	56.9	397	AC1574	acetate kinase hom
3	41	56.9	397	AC1574	acetate kinase hom
4	40	55.6	383	2	hypothetical prote
5	39	54.2	409	2	Kruppel-like prote
6	39	54.2	1117	2	hypothetical prote
7	38	52.8	378	2	membrane-associate
8	38	52.8	469	2	NAC2-like protein
9	37	51.4	319	2	lytB protein (impo
10	37	51.4	421	2	XPMC2 protein - Af
11	37	51.4	433	2	pectinesterase (EC
12	37	51.4	988	1	protein kinase C (
13	37	51.4	1174	2	protein kinase C (
14	37	51.4	1294	2	probable membrane
15	36.5	50.7	230	2	hypothetical prote
16	36	50.0	35	2	T-cell receptor be
17	36	50.0	282	2	hypothetical prote
18	36	50.0	372	2	hypothetical prote
19	36	50.0	419	2	GTP-binding regula
20	36	50.0	442	2	probable RNA helic
21	36	50.0	483	2	catalase (EC 1.11.
22	36	50.0	693	2	C30A5.5 protein -
23	36	50.0	814	2	probable allene ox
24	36	50.0	944	2	hypothetical prote
25	36	50.0	1226	2	hypothetical prote
26	35.5	49.3	224	2	MG243 homolog H91
27	35.5	49.3	266	2	germin homolog F21
28	35	48.6	47	2	viral capsid prote
29	35	48.6	62	2	ORF MSV185 hypothe

30 35 48.6 66 2 A27406 toxin CasII - Mexi  
31 35 48.6 80 2 PC4272 cdu3 protein - Clo  
32 35 48.6 155 2 D82152 conserved hypothet  
33 35 48.6 182 2 H71899 hypothetical prote  
34 35 48.6 207 2 AF0199 probable lipoprote  
35 35 48.6 218 2 T33545 hypothetical prote  
36 35 48.6 228 2 S67199 hypothetical prote  
37 35 48.6 238 2 S77699 inner cell wall ma  
38 35 48.6 265 2 AF2007 hypothetical prote  
39 35 48.6 276 2 S27641 hypothetical prote  
40 35 48.6 308 2 AF1990 hypothetical prote  
41 35 48.6 375 2 H82329 alanine racemase,  
42 35 48.6 393 2 E64239 acetate kinase (EC  
43 35 48.6 396 2 D81339 acetate kinase (EC  
44 35 48.6 415 2 A11611 competence-damage  
45 35 48.6 483 1 T43234 protein kinase (EC

## ALIGNMENTS

### RESULT 1

A39035

ribonuclease-related anti-tumor protein - northern leopard frog (fragment)

C:Species: Rana pipiens (northern leopard frog)

C>Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 30-Jun-1993

C:Accession: A39035

R:Ardelt, W.; Mikulski, S.M.; Shogen, K.

J. Biol. Chem. 266, 245-251, 1991

A:Title: Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and earl

A:Reference number: A39035; MUID:91093131; PMID:1985896

A:Accession: A39035

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-104 <ARD>

C:Superfamily: pancreatic ribonuclease

Query Match 59.7%; Score 43; DB 2; Length 104;

Best Local Similarity 61.5%; Pred. No. 1.7;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNTFC 13

Db 75 CKYKUKKSTNKC 87

### RESULT 2

AC1574

acetate kinase homolog AckA2 [imported] - Listeria innocua (strain Clp11262)

C:Species: Listeria innocua

C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C:Accession: AC1574

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Psihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wenland,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1574

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96363.1; PID:gi61413591; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: AckA2

C:Superfamily: acetate kinase

Query Match 56.9%; Score 41; DB 2; Length 397;

Best Local Similarity 72.7%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11  
 |||||  
 Db 214 CAIEAGKSVNT 224

## RESULT 3

acacate kinase homolog AckA2 [imported] - Listeria monocytogenes (strain EGD-e)  
 C;Species: Listeria monocytogenes  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C;Accession: AH1220  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A;Title: Comparative genomics of Listeria species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AH1220  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-397 <GLA>  
 A;Cross-references: GB:NC\_003210; PIDN:CAC99246.1; PID:g16410584; GSPDB:GN00177  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: AckA2  
 C;Superfamily: acetate kinase

Query Match 56.9%; Score 41; DB 2; Length 397;  
 Best Local Similarity 72.7%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 1 CASELGKSTNT 11  
 |||||  
 Db 214 CAIEAGKSVNT 224

## RESULT 4

hypothetical protein A231L - Chlorella virus PBCV-1  
 C;Species: Chlorella virus PBCV-1  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T17722  
 R;Graves, M.V.; Van Etten, J.L.  
 submitted to the EMBL Data Library, May 1999  
 A;Reference number: Z18806  
 A;Accession: T17722  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-383 <GRA>  
 A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96599.1  
 A;Experimental source: specific host Chlorella strain NC64A  
 C;Genetics:  
 A;Note: A231L

Query Match 55.6%; Score 40; DB 2; Length 383;  
 Best Local Similarity 46.2%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 6; Indels 6; Gaps 0;

QY 1 CASELGKSTNTFC 13  
 |||||  
 Db 304 CVMNIGKKNEFC 316

## RESULT 5

S42384  
 Kruppel-like protein - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 07-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Sep-1997  
 C;Accession: S42384  
 R;Smith, A.

submitted to the EMBL Data Library, March 1994  
 A;Reference number: S42368  
 A;Accession: S42384  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-409 <SM1>  
 A;Cross-references: EMBL:Z30662; NID:g459581; PID:g459587  
 C;Genetics:  
 A;Introns: 10/3; 48/3; 182/3; 315/3; 360/3

Query Match 54.2%; Score 39; DB 2; Length 409;  
 Best Local Similarity 58.3%; Pred. No. 33;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNTF 12  
 |||||  
 Db 229 CKSDLGSPQTQF 240

## RESULT 6

C85018  
 hypothetical protein AT4g01400 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C;Accession: C85018  
 R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
 Nature 402, 769-777, 1999  
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A;Reference number: A85001; MUID:20083488; PMID:10617198  
 A;Accession: C85018  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1117 <STO>  
 A;Cross-references: GB:NC\_001268; NID:g7267637; PIDN:CAB80949.1; GSPDB:GN00140  
 C;Genetics:  
 A;Gene: AT4g01400  
 A;Map position: 4

Query Match 54.2%; Score 39; DB 2; Length 1117;  
 Best Local Similarity 58.3%; Pred. No. 85;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CASELGKSTNTF 12  
 |||||  
 Db 922 CLSELGELSSTF 933

## RESULT 7

G97177  
 membrane-associated sensory histidine kinase-like ATPase [imported] - Clostridium acetol  
 C;Species: Clostridium acetobutylicum  
 C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C;Accession: G97177  
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
 A;Reference number: A96900; MUID:21359325; PMID:21359325  
 A;Accession: G97177  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-378 <KUR>  
 A;Cross-references: GB:AE001437; PIDN:AAK80210.1; PID:g15025255; GSPDB:GN00168  
 A;Experimental source: Clostridium acetobutylicum ATCC824  
 C;Genetics:  
 A;Gene: CAC2253

Query Match 52.8%; Score 38; DB 2; Length 378;  
 Best Local Similarity 63.6%; Pred. No. 47;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11  
 |||||

Db 278 CASDIEKSTTT 288

# RESULT 8

T46230

NAC2-like protein - Arabidopsis thaliana  
 N;Alternate names: protein T9C5.120  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C;Accession: T46230  
 R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the Protein Sequence Database, December 1999  
 A;Reference number: Z23026  
 A;Accession: T46230  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-469 <RIE>  
 A;Cross-references: EMBL:AL132964  
 A;Experimental source: cultivar Columbia; BAC clone T9C5  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 60/1; 153/3; 233/3; 332/3; 402/3  
 A;Note: T9C5.120

Query Match 52.8%; Score 38; DB 2; Length 469;  
 Best Local Similarity 80.0%; Pred. No. 57;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 SELGKSTNTF 12

|||||

Db 260 SELGSSYNF 269

# RESULT 9

A84947

lytB protein [imported] - Buchnera sp. (strain APS)  
 C;Species: Buchnera sp.  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C;Accession: A84947  
 R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Nature 407, 81-86, 2000  
 A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.  
 A;Reference number: A84930; MUID:20445173; PMID:10993077  
 A;Accession: A84947  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-319 <STO>  
 A;Cross-references: GB:AP000398; GSPDB:GN00144  
 A;Experimental source: strain APS  
 C;Genetics:  
 A;Gene: lytB; BU147  
 C;Superfamily: penicillin tolerance protein

Query Match 51.4%; Score 37; DB 2; Length 319;  
 Best Local Similarity 70.0%; Pred. No. 61;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SELGKSTNTF 12

|||||

Db 232 AELGKETGF 241

# RESULT 10

S53818

XPM22 protein - African clawed frog  
 C;Species: Xenopus laevis (African clawed frog)  
 C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 21-Jul-2000  
 C;Accession: S53818  
 R;Su, J.Y.; Maller, J.L. Mol. Gen. Cloning 246, 387-396, 1995  
 A;Title: Cloning and expression of a Xenopus gene that prevents mitotic catastrophe in fibroblasts  
 A;Reference number: S53818; MUID:95157530; PMID:7854324  
 A;Accession: S53818

A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-421 <SUU>  
 A;Cross-references: GB:U10185; NID:G595379; PIDN:AAA82179.1; PID:G595380

Query Match 51.4%; Score 37; DB 2; Length 421;  
 Best Local Similarity 54.5%; Pred. No. 79;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELCKSTNT 11

|||||

Db 13 CSGSLGTANT 23

# RESULT 11

S70914

pectinesterase (EC 3.1.1.11) precursor - Erwinia chrysanthemi  
 N;Alternate names: pectin methyltransferase  
 C;Species: Erwinia chrysanthemi  
 C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000  
 C;Accession: S70914; S70913  
 R;Shevchik, V.E.; Condemine, G.; Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J. submitted to the EMBL Data Library, February 1995  
 A;Description: Characterization of pectin methyltransferase B, an outer membrane lipoprotein  
 A;Reference number: S70914  
 A;Accession: S70914  
 A;Molecule type: DNA  
 A;Residues: 1-433 <SHE>  
 A;Cross-references: EMBL:X84665; NID:G1212890; PIDN:CAA59151.1; PID:G1212891  
 R;Shevchik, V.E.; Condemine, G.; Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J. Mol. Microbiol. 19, 455-466, 1996  
 A;Title: Characterization of pectin methyltransferase B, an outer membrane lipoprotein  
 A;Reference number: S70913; MUID:96228695; PMID:8830237  
 A;Accession: S70913  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 1-53;234-361 <SHA>  
 A;Cross-references: EMBL:X84665; NID:G1212890; PIDN:CAA59151.1  
 A;Experimental source: strain 3937  
 C;Genetics:  
 A;Gene: pEMB  
 A;Start codon: GTG  
 C;Superfamily: pectinesterase pEMB  
 C;Keywords: carboxylic ester hydrolase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-433/Product: pectinesterase #status predicted <MAT>

Query Match 51.4%; Score 37; DB 2; Length 433;  
 Best Local Similarity 46.2%; Pred. No. 81;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELCKSTNTFC 13

|||||

Db 192 CATKAGATINTTC 204

# RESULT 12

S35362

protein kinase C (EC 2.7.1.1-) pck1 - fission yeast (Schizosaccharomyces pombe)  
 N;Alternate names: pombe C-kinase 1  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 06-Oct-2000  
 C;Accession: S35362; T37866; T38203  
 R;Toda, T.; Shimanuki, M.; Yanagida, M. EMBO J. 12, 1987-1995, 1993  
 A;Title: Two novel protein kinase C-related genes of fission yeast are essential for cell growth  
 A;Reference number: S35362; MUID:93259141; PMID:8491190  
 A;Accession: S35362  
 A;Molecule type: DNA  
 A;Residues: 1-988 <TOD>  
 A;Cross-references: EMBL:D14337; NID:G303938; PIDN:BAA03267.1; PID:G303939  
 R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V. submitted to the EMBL Data Library, February 1996

A:Reference number: Z21750  
 A:Accession: T37866  
 A:Molecule type: DNA  
 A:Residues: 93-988 <DEV>  
 A:Cross-references: EMBL:Z69795; PIDN:CAA93697.1; GSPDB:GN00066; SPDB:SPAC17G8.14c  
 F:Grishchuk, K.; McIntosh, J.R.; Devlin, K.; Churcher, C.; Barrell, B.G.; Rajandream, M.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z21778  
 A:Accession: T38203  
 A:Molecule type: DNA  
 A:Residues: 1-26, 'AVASTKNP', 36-137 <ORI>  
 A:Cross-references: EMBL:Z69730; PIDN:CAA93602.1; GSPDB:GN00066; SPDB:SPAC22H10.01c  
 A:Experimental source: strain 972h-; cosmid c22H10  
 C:Genetics:  
 A:Gene: pck1  
 A:Map position: 1L  
 A:Introns: 767/2; 834/3; 897/3; 959/1  
 C:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology; p  
 C:Keywords: ATP; duplication; phospholipid binding; phosphotransferase; serine/threonine  
 F:414-461/Domain: protein kinase C zinc-binding repeat homology <KZ1>  
 F:481-530/Domain: protein kinase C zinc-binding repeat homology <KZ2>  
 F:682-923/Domain: protein kinase C zinc-binding repeat homology <KIN>  
 F:670-678/Domain: protein kinase ATP-binding motif

Query Match 51.4%; Score 37; DB 1; Length 988;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LGKSTNTFC 13

Db 817 LGNTTSTFC 825

# RESULT 13

T43051  
 protein kinase C (EC 2.7.1.-) PKC1 - fungus (Cochliobolus heterostrophus)  
 C:Species: Cochliobolus heterostrophus, Bipolaris maydis  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T43051  
 R:Oeser, B.M.  
 FEMS Microbiol. Lett. 165, 273-280, 1998  
 A:Title: PKC1, encoding a protein kinase C, and FAT1, encoding a fatty acid transporter  
 A:Reference number: Z22303; MUID:98415124; PMID:9742699  
 A:Accession: T43051  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1174 <ORS>  
 A:Cross-references: EMBL:Y15839; NID:G2687848; PIDN:CAA75801.1; PID:G2687849  
 A:Experimental source: strain C2 (ATCC 48329)  
 C:Genetics:  
 A:Gene: pck1  
 A:Introns: 146/1; 234/1; 559/1; 952/2; 1019/3; 1155/1  
 C:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology; p  
 C:Keywords: ATP; phospholipid binding; phosphotransferase; serine/threonine-specific pro  
 F:459-506/Domain: protein kinase C zinc-binding repeat homology <KZN1>  
 F:527-576/Domain: protein kinase C zinc-binding repeat homology <KZN2>

Query Match 51.4%; Score 37; DB 2; Length 1174;  
 Best Local Similarity 46.7%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

Qy 1 CASEL--GKSTNTFC 13

Db 996 CKERMWYGSITSTFC 1010

# RESULT 14

S77690  
 probable membrane protein YOL075c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein O1125; hypothetical protein O1130; hypothetical  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 19-Apr-2002  
 C:Accession: S77690; S66767; S66768

R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S66756  
 A:Accession: S77690  
 A:Molecule type: DNA  
 A:Residues: 1-1294 <ALE>  
 A:Cross-references: EMBL:Z74816; MIPS:YOL075c  
 A:Note: this is a revision to the sequence from reference S66756  
 A:Accession: S66767  
 A:Molecule type: DNA  
 A:Residues: 1-179, 'TTRTGVLVVKRED' <ALW>  
 A:Cross-references: EMBL:Z74816  
 A:Experimental source: strain S288C  
 A:Note: this sequence has been revised in reference S77690  
 A:Note: this was assumed to be protein YOL074c  
 A:Accession: S66768  
 A:Molecule type: DNA  
 A:Residues: 200-1294 <ALP>  
 A:Cross-references: EMBL:Z74817  
 A:Experimental source: strain S288C  
 A:Note: this sequence has been revised in reference S77690  
 A:Note: this was assumed to be the complete sequence of protein YOL075c  
 C:Genetics:  
 A:Cross-references: SGD:S0005435  
 A:Map position: 15L  
 A:Note: YOL075c  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
 F:45-263/Domain: ATP-binding cassette homology <ABC1>  
 F:62-69/Region: nucleotide-binding motif A (P-loop)  
 F:376-392/Domain: transmembrane #status predicted <TM1>  
 F:469-485/Domain: transmembrane #status predicted <TM2>  
 F:496-512/Domain: transmembrane #status predicted <TM3>  
 F:606-622/Domain: transmembrane #status predicted <TM4>  
 F:710-916/Domain: ATP-binding cassette homology <ABC2>  
 F:727-734/Region: nucleotide-binding motif A (P-loop)  
 F:1042-1058/Domain: transmembrane #status predicted <TM5>  
 F:1125-1141/Domain: transmembrane #status predicted <TM6>  
 F:1177-1193/Domain: transmembrane #status predicted <TM7>  
 F:1269-1285/Domain: transmembrane #status predicted <TM8>

Query Match 51.4%; Score 37; DB 2; Length 1294;  
 Best Local Similarity 58.3%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CASELGRKSTNTFC 12

Db 1161 CGERLGIMTNTFC 1172

# RESULT 15

## S09778

hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)  
 C:Species: human cytomegalovirus, human herpesvirus 5  
 A:Note: host Homo sapiens (man)  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Oct-1999  
 C:Accession: S09778

R:Cheer, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;  
 M.; Barrell, B.G.  
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
 A:Reference number: S09749; MUID:90269039; PMID:2161319  
 A:Accession: S09778  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-230 <CHE>

A:Cross-references: EMBL:X17403; NID:G959591; PIDN:CAA35448.1; PID:G959620  
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
 C:Superfamily: human cytomegalovirus hypothetical protein UL16  
 C:Keywords: glycoprotein; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-230/Product: hypothetical protein UL16 #status predicted <MAT>  
 F:189-206/Domain: transmembrane #status predicted <TMM>

F:35,41,68,84,95,101,132,145/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 50.7%; Score 36.5; DB 2; Length 230;

Best Local Similarity 57.1%; Pred. No. 55;

Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 CASELG-KSTNTFC 13

||:|||||:|

Db 25 CAVDLGSKSNSTC 38

Search completed: September 5, 2004, 10:01:14

Job time : 11.9798 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 09:38:39 ; Search time 5.90909 Seconds  
(without alignments)  
114.554 Million cell updates/sec

Title: US-09-761-636A-5

Perfect score: 72

Sequence: 1 CASELCKSTNTFC 13

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	75.0	354	1	VEGD HUMAN
2	51	70.8	326	1	VEGD RAT
3	51	70.8	358	1	VEGD MOUSE
4	43	59.7	104	1	RN30_FANPI
5	41	56.9	397	1	ACK2_LISIN
6	41	56.9	397	1	ACK2_LISMO
7	39	54.2	451	1	YNV5_CAEEL
8	37	51.4	169	1	11F9_HUMAN
9	37	51.4	319	1	ISPH_BUCAI
10	37	51.4	407	1	RPF2_HUMAN
11	37	51.4	433	1	PNEB_ERWCH
12	37	51.4	988	1	PKK1_SCHPO
13	37	51.4	1174	1	KPC1_COCHE
14	37	51.4	1294	1	YOH5_YEAST
15	36.5	50.7	230	1	UL16_HCMVA
16	36	50.0	282	1	PNK_HALNI
17	36	50.0	362	1	XR86_CAEEL
18	36	50.0	483	1	BCA_STRVL
19	36	50.0	673	1	ESR2_MICUN
20	36	50.0	944	1	TPSK_SCHPO
21	36	50.0	1360	1	CTNG_XENLA
22	35.5	49.3	224	1	Y243_MYCPN
23	35.5	49.3	2156	1	RPI_HUMAN
24	35	48.6	66	1	SCX2_CENSU
25	35	48.6	238	1	CW14_YEAST
26	35	48.6	313	1	ISPH_BUCAI
27	35	48.6	361	1	ALR1_VIBCH
28	35	48.6	393	1	ACKA_MYCGE
29	35	48.6	396	1	ACKA_CAMUE
30	35	48.6	576	1	FTS1_BUCBP
31	35	48.6	602	1	TRAN_ECOLI
32	35	48.6	758	1	PARC_RHIME
33	35	48.6	855	1	ENV_FVIA2

RESULT 1				
VEGD_HUMAN	ID	VEGD_HUMAN	STANDARD;	PRT; 354 AA.
AC	043915;			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).			
DE	Growth factor) (FIGF).			
GN	FIGF OR VEGFD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;			
RC	MEDLINE=97349118; PubMed=9205122;			
RX	Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;			
RA	"Molecular cloning of a novel vascular endothelial growth factor, VEGF-D";			
RT	Genomics 42:483-488 (1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;			
RC	MEDLINE=98140120; PubMed=9479493;			
RX	Rocchigiani M., Lestingi M., Luddi A., Orlandini M., Franco B., Rossi E., Ballabio A., Zuffardi O., Oliviero S.;			
RA	"Human FIGF: Cloning, gene structure, and mapping to chromosome Xp22.1 between the FIGA and the GRPR genes.";			
RT	Genomics 47:207-216 (1998).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=98118549; PubMed=9435229;			
RX	Achen M.G., Jeltsch M., Kukk E., Maekinen T., Vitali A., Wilks A.F., Alitalo K., Stacker S.A.;			
RA	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";			
RT	Proc. Natl. Acad. Sci. U.S.A. 95:548-553 (1998).			
RL	[4]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;			
RC	MEDLINE=22388257; PubMed=12477932;			
RX	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			

Q09349 caenorhabdi  
O13356 cryptococcu  
P24583 saccharomyc  
P09831 escherichia  
O60014 kluyveromyc  
Q91206 mus musculu  
Q12522 saccharomyc  
P32572 saccharomyc  
O31605 bacillus su  
P57628 buchnera ap  
Q29100 sus scrofa  
Q91574 xenopus lae

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP PROCESSING AND SEQUENCE OF 89-94; 100-105 AND 206-213.  
 RX MEDLINE-20011413; PubMed-10542248;  
 RA Stacker S.A., Stenvers K.L., Caesar C., Vitali A., Domagala T.,  
 RA Nice E.C., Roufail S., Simpson R.J., Moritz R., Karpanen T.,  
 RA Alitalo K., Achen M.G.;  
 RT "Biosynthesis of vascular endothelial growth factor-D involves  
 proteolytic processing which generates non-covalent homodimers.";  
 RL J. Biol. Chem. 274:32127-32136(1999).  
 CC -!- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis  
 and endothelial cell growth, stimulating their proliferation and  
 migration and also has effects on the permeability of blood  
 vessels. May function in the formation of the venous and lymphatic  
 vascular systems during embryogenesis, and also in the maintenance  
 of differentiated lymphatic endothelium in adults. Binds and  
 activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.  
 CC -!- SUBUNIT: Homodimer; non-covalent and antiparallel.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Highly expressed in lung, heart, small  
 intestine and fetal lung, and at lower levels in skeletal muscle,  
 colon, and pancreas.  
 CC -!- PTM: Undergoes a complex proteolytic maturation which generates a  
 variety of processed secreted forms with increased activity toward  
 VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer  
 linked by disulfide bonds before secretion. The fully processed  
 VEGF-D is composed mostly of two VEGF homology domains (VHDs)  
 bound by non-covalent interactions.  
 CC -!- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D89630; BAA24264.1; -  
 DR EMBL; Y12863; CAA73370.1; -  
 DR EMBL; Y12864; CAA73371.1; -  
 DR EMBL; Y12865; CAA73371.1; JOINED.  
 DR EMBL; Y12866; CAA73371.1; JOINED.  
 DR EMBL; Y12867; CAA73371.1; JOINED.  
 DR EMBL; Y12868; CAA73371.1; JOINED.  
 DR EMBL; Y12869; CAA73371.1; JOINED.  
 DR EMBL; Y12870; CAA73371.1; JOINED.  
 DR EMBL; AJ000185; CAA03942.1; -  
 DR EMBL; BC027948; AAB27948.1; -  
 DR HSSP; P15692; 1VPP.  
 DR MIM; 300091; -  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0005161; F:platelet-derived growth factor receptor bin. .; TAS.  
 DR GO; GO:0005102; F:receptor binding; TAS.  
 DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.  
 DR InterPro; IPR004153; CXXC repeat.  
 DR InterPro; IPR000072; PD\_growth\_factor.  
 DR Pfam; PF03128; CXXC; 3.  
 DR Pfam; PF0341; PDGF; 1.  
 DR Prodom; PD001629; PD\_growth\_factor; 1.  
 DR SMART; SM00141; PDGF; 1.  
 DR PROSITE; PS00249; PDGF 1; 1.  
 DR PROSITE; PS0278; PDGF 2; 1.  
 KW Angiogenesis; Mitogen; Growth factor; Glycoprotein; Signal; Repeat;  
 KW Cleavage on pair of basic residues; Multigene family.  
 FT SIGNAL 1 21 POTENTIAL.

FT PROPEP 22 88  
 FT CHAIN 89 205  
 FT PROPEP 206 354  
 FT DOMAIN 222 318  
 FT REPEAT 222 237  
 FT REPEAT 258 273  
 FT REPEAT 277 293  
 FT REPEAT 301 318  
 FT DISULFID 111 153  
 FT DISULFID 142 189  
 FT DISULFID 146 191  
 FT DISULFID 136 136  
 FT DISULFID 145 145  
 FT CARBOHYD 155 155  
 FT CARBOHYD 185 185  
 FT CARBOHYD 287 287  
 SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;  
 Query Match 75.0%; Score 54; DB 1; Length 354;  
 Best Local Similarity 100.0%; Pred.No. 0.018; Mismatches 0; Indels 0; Gaps 0;  
 Matches 11; Conservative 0;  
 QY 2 ASELGKSTNTF 12  
 DB 121 ASELGKSTNTF 131  
 |||||  
 |||||  
 RESULT 2  
 VEGD RAT STANDARD; PRT; 326 AA.  
 AC Q35251;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced  
 growth factor) (FIGF).  
 GN FIGF OR VEGFD.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID:10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley;  
 RA Yamada Y., Hirata Y., Nezu J., Shimane M.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis  
 and endothelial cell growth, stimulating their proliferation and  
 migration and also has effects on the permeability of blood  
 vessels. May function in the formation of the venous and lymphatic  
 vascular systems during embryogenesis, and also in the maintenance  
 of differentiated lymphatic endothelium in adults. Binds and  
 activates VEGFR-3 (Flt4) receptor (By similarity).  
 CC -!- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- PTM: Undergoes a complex proteolytic maturation which generates a  
 variety of processed secreted forms with increased activity toward  
 VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer  
 linked by disulfide bonds before secretion. The fully processed  
 VEGF-D is composed mostly of two VEGF homology domains (VHDs)  
 bound by non-covalent interactions (By similarity).  
 CC -!- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF014827; AAB66557.1; -

```

DR HSSP; PL5692; 1VPP.
DR InterPro; IPR004153; CXXC_repeat.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF03128; CXXC; 1.
DR Pfam; PF03341; PDGF; 1.
DR ProDom; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS0278; PDGF_2; 1.
KW Angiogenesis; Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 326
FT DOMAIN 227 317
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 306 317
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 70.8%; Score 51; DB 1; Length 326;
Best Local Similarity 90.9%; Pred. No. 0.06;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTF 12
DB 126 ASELGKSTNTF 136
|||||:|||||
|||||:|||||

RESULT 3
VEGD MOUSE
ID VEGD MOUSE STANDARD; PRT; 358 AA.
AC P97926;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
DE growth factor) (FIGF).
GN FIGF OR VEGFD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Fibroblast;
RX MEDLINE=97030254; PubMed=8876195;
RA Orlandini M., Marconcin L., Ferruzzi R., Oliviero S.;
RT "Identification of a c-fos-induced gene that is related to the
RT platelet-derived growth factor/vascular endothelial growth factor
RT family".
RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D".
RL Genomics 42:483-488(1997).
RN [3]
RP DEVELOPMENTAL STAGE.

```

RX MEDLINE=98288130; PubMed=9622638;  
 RA Avantiato V., Orlandini M., Acampora D., Oliviero S., Simeone A.;  
 RT "Embryonic expression pattern of the murine figf gene, a growth factor  
 RT belonging to platelet-derived growth factor/vascular endothelial  
 RT growth factor family".  
 RL Mech. Dev. 73:221-224(1998).  
 RN [4]  
 RP RECEPTOR SPECIFICITY  
 RX MEDLINE=21276411; PubMed=11279005;  
 RA Baldwin M.E., Carimel B., Nice E.C., Roufail S., Hall N.E.,  
 RA Stenvers K.L., Karkkainen M.J., Alitalo K., Stacker S.A., Achen M.G.;  
 RT "The specificity of receptor binding by vascular endothelial growth  
 RT factor-d is different in mouse and man".  
 RL J. Biol. Chem. 276:19166-19171(2001).  
 CC -!- FUNCTION: Growth factor active in angiogenesis. Lymphangiogenesis  
 CC and endothelial cell growth, stimulating their proliferation and  
 CC migration and also has effects on the permeability of blood  
 CC vessels. May function in the formation of the venous and lymphatic  
 CC vascular systems during embryogenesis, and also in the maintenance  
 CC of differentiated lymphatic endothelium in adults. Binds and  
 CC activates VEGFR-3 (Flt4) receptor.  
 CC -!- SUBUNIT: Homodimer; non-covalent and antiparallel.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.  
 CC -!- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several  
 CC body structures and organs of the embryo such as limb buds,  
 CC acoustic ganglion, teeth, heart, anterior pituitary as well as  
 CC lung and kidney mesenchyme, liver, derma, and perosteum of the  
 CC vertebral column.  
 CC -!- INDUCTION: By the transcription factor c-fos.  
 CC -!- PTM: Undergoes a complex proteolytic maturation which generates a  
 CC variety of processed secreted forms with increased activity toward  
 CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer  
 CC linked by disulfide bonds before secretion. The fully processed  
 CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)  
 CC bound by non-covalent interactions (By similarity).  
 CC -!- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X99572; CAA67892.1; -.  
 CC EMBL; D89628; BAAL4002.1; -.  
 CC HSSP; PL5692; 1VPP.  
 CC PMMA-2DPAGE; P97946; -.  
 CC MGD; MGI:108037; Figf.  
 CC GO; GO:000576; C:extracellular; IDA.  
 CC GO; GO:0008083; F:growth factor activity; IDA.  
 CC GO; GO:0005515; F:protein binding; IPT.  
 CC GO; GO:0008283; P:cell proliferation; IDA.  
 CC InterPro; IPR004153; CXXC\_repeat.  
 CC InterPro; IPR000072; PD\_growth\_factor.  
 CC Pfam; PF03128; CXXC; 2.  
 CC Pfam; PF03341; PDGF; 1.  
 CC ProDom; PD001629; PD\_growth\_factor; 1.  
 CC SMART; SM00141; PDGF; 1.  
 CC PROSITE; PS00249; PDGF\_1; 1.  
 CC PROSITE; PS0278; PDGF\_2; 1.  
 KW Angiogenesis; Mitogen; Growth factor; Glycoprotein; Signal; Repeat;  
 KW Cleavage on pair of basic residues; Multigene family.  
 FT SIGNAL 1 21  
 FT PROPEP 22 93  
 FT CHAIN 94 210  
 FT PROPEP 211 358  
 FT DOMAIN 227 323  
 FT REPEAT 227 242  
 FT REPEAT 263 278

```

FT REPEAT 282 298 3.
FT REPEAT 306 323 4.
FT DISULFID 116 158 BY SIMILARITY.
FT DISULFID 147 194 BY SIMILARITY.
FT DISULFID 151 196 BY SIMILARITY.
FT DISULFID 141 141 INTERCHAIN (BY SIMILARITY).
FT DISULFID 150 150 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 358 AA; 40908 MW; 6636BL7PBF07037C CRC64;

Query Match 70.8%; Score 51; DB 1; Length 358;
Best Local Similarity 90.9%; Pred. No. 0.066;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ASELGKSTNTP 12
Db 126 ASELGKSTNTP 136
|||||:|||||

RESULT 4
RN30_RANPI STANDARD; PRT; 104 AA.
ID RN30_RANPI
AC P22069;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE P-30 protein (EC 3.1.27.-) (Onconase).
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN NCBI_TaxID=8404;
RP SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=91093131; PubMed=1985896;
RA Ardel W., Mikulski S.M., Shogen K.;
RT "Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases.";
RL J. Biol. Chem. 266:245-251(1991).
RN [2]

3D-STRUCTURE MODELING.
RP MEDLINE=93066156; PubMed=1438177;
RX Mosimann S.C., Johns K.L., Ardel W., Mikulski S.M., Shogen K.;
RA James M.N.G.;
RT "Comparative molecular modeling and crystallization of P-30 protein: a novel antitumor protein of Rana pipiens oocytes and early embryos.";
RL Proteins 14:392-400(1992).
RN [3]

X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RP MEDLINE=94166079; PubMed=8120892;
RX Mosimann S.C., Ardel W., James M.N.G.;
RT "Refined 1.7 A x-ray crystallographic structure of P-30 protein, an amphibian ribonuclease with anti-tumor activity.";
RL J. Mol. Biol. 236:1141-1153(1994).
RN [4]

FUNCTION: Basic protein with antiproliferative/cytotoxic activity against several tumor cell lines in vitro, as well as antitumor in vivo. It exhibits a ribonuclease-like activity against high molecular weight ribosomal RNA.
CC -1- DEVELOPMENTAL STAGE: Early embryos (up to four blastomere stage).
CC -1- SIMILARITY: Belongs to the pancreatic ribonuclease family.
DR PDB; LONG; 31-JAN-94.
DR InterPro; IPR001427; RNaseA.
DR Pfam; PF000074; RNaseA; 1.
DR ProDom; PD000535; RNaseA; 1.
DR SMART; SM00092; RNaseA; PC; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; 1.
DR HydroLase; Nuclease; Endonuclease; 3D-structure;
KW Pyrrolidone carboxylic acid. PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1
FT ACT_SITE 10 10

```

```

FT ACT_SITE 31 31
FT ACT_SITE 97 97
FT DISULFID 19 68
FT DISULFID 30 75
FT DISULFID 48 90
FT DISULFID 87 104
FT HELIX 3 10
FT STRAND 11 12
FT HELIX 19 22
FT TURN 23 24
FT TURN 26 30
FT STRAND 33 38
FT HELIX 41 45
FT HELIX 46 48
FT TURN 49 50
FT STRAND 55 58
FT STRAND 63 70
FT TURN 74 75
FT STRAND 77 84
FT STRAND 86 91
FT TURN 92 93
FT STRAND 94 101
SQ SEQUENCE 104 AA; 11845 MW; 22A753C2F9E566B4 CRC64;

Query Match 59.7%; Score 43; DB 1; Length 104;
Best Local Similarity 61.5%; Pred. No. 0.57;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CASELGKSTNTPC 13
Db 75 CKYKLGKSTNPKFC 87
|:|:|:|:|:|

RESULT 5
ACK2 LISIN STANDARD; PRT; 397 AA.
ID ACK2 LISIN
AC Q92CN9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetate Kinase 2 (EC 2.7.2.1) (Acetokinase 2).
GN ACKA2 OR LIN1132.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]

SEQUENCE FROM N.A.
RP STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -1- CATALYTIC ACTIVITY: ATP + acetate = ADP + acetyl phosphate.
CC -1- PATHWAY: Conversion of acetate to acetyl-CoA; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the acetokinase family.
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
-----
CC EMBL; AL596167; CAC96363.1; --
CC PIR; AC1574; AC1574
CC ListList; LIN01132; --
CC HAMAP; MF_00020; -- 1.
CC InterPro; IPR004372; AcKa.
CC Pfam; PF00871; Acetate_kinase; 1.
CC PRINTS; PR00471; ACETATEKINASE.
CC TIGRFAMs; TIGR00016; acKa; 1.
CC PROSITE; PS01075; ACETATE_KINASE_1; 1.
CC TRANSFERASE; Kinase; Complete proteome.
SQ SEQUENCE 397 AA; 43115 MW; 5922544EF92CBF51 CRC64;

Query Match 56.9%; Score 41; DB 1; Length 397;
Best Local Similarity 72.7%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11
DB 214 CAIEAGKSVNT 224

RESULT 6
ACK2 LISMO
ID ACK2 LISMO STANDARD; PRT; 397 AA.
AC ORY7V;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetate kinase 2 (EC 2.7.2.1) (Acetokinase 2).
GN ACK2 OR LM01168.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madsen E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + acetate = ADP + acetyl phosphate.
CC -!- PATHWAY: Conversion of acetate to acetyl-CoA; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the acetokinase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AL591978; CAC99246.1; --
CC PIR; AH1220; AH1220.
CC ListList; LM001168; --
CC HAMAP; MF_00020; -- 1.
CC InterPro; IPR000890; Acetate_kin.

```

```

DR Pfam; PF00871; Acetate kinase; 1.
DR PRINTS; PR00471; ACETATEKINASE.
DR TIGRFAMs; TIGR00016; acKa; 1.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
DR PROSITE; PS01076; ACETATE_KINASE_2; 1.
DR TRANSFERASE; Kinase; Complete proteome.
SQ SEQUENCE 397 AA; 43132 MW; 6E295A5A5PD5C5B CRC64;

Query Match 56.9%; Score 41; DB 1; Length 397;
Best Local Similarity 72.7%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11
DB 214 CAIEAGKSVNT 224

RESULT 7
YNV5 CAEEL
ID YNV5 CAEEL STANDARD; PRT; 451 AA.
AC P34588;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein T16H12.5 in chromosome III.
GN T16H12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 MATH domain.
CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; Z30662; CAA83138.2; --
CC WormPep; T16H12.5; CE29054.
CC InterPro; IPR000210; BTB_POZ.
CC InterPro; IPR002083; MATH.
CC InterPro; IPR008974; Traf_dom.
CC Pfam; PF00651; BTB; 1.
CC Pfam; PF00917; MATH; 1.
CC SMART; SM00225; BTB; 1.
CC SMART; SM00061; MATH; 1.
CC PROSITE; PS00097; BTB; 1.
CC PROSITE; PS0144; MATH; 1.
KW Hypothetical protein.
FT DOMAIN 95 225 MATH.
FT DOMAIN 265 338 BTB.
SQ SEQUENCE 451 AA; 51062 MW; B36B1C618FBB3A3 CRC64;

Query Match 54.2%; Score 39; DB 1; Length 451;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNTF 12
DB 271 CAXSDLGSTQTF 282

```

```

DR MIM; 605542; -.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0009613; P:response to pest/pathogen/parasite; TAS.
DR InterPro; IPR008996; Cytok_IL1-like.
DR InterPro; IPR000975; Interleukin_1.
DR Pfam; PF00340; IL1; 1.
DR PRODOM; PD002536; Interleukin_1; 1.
DR SMART; SM0125; IL1; 1.
DR PROSITE; PS00253; INTERLEUKIN_1; FALSE_NEG.
DR Cytokine; Multigene family.
SQ SEQUENCE 169 AA; 18721 MW; F00A9243705F4154 CRC64;

Query Match 51.4%; Score 37; DB 1; Length 169;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SELGKSTNT 11
DB 153 SELGKSYNT 161

RESULT 9
ISPH_BUCAI
ID ISPH_BUCAI STANDARD; PRT; 319 AA.
AC P57247;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
GN ISPH OR LYTB OR BUI47.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum)
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; PubMed=109931077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT Buchnera sp. APS.;
RL Nature 407:81-86(2000).
CC -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
CC into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate
CC (DMAPP) (By similarity).
CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
CC (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
CC (last) step.
CC -!- SIMILARITY: Belongs to the isph family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001118; BAB12865.1; -.
CC HAMAP; MF 00191; -.
CC InterPro; IPR003451; LytB.
CC Pfam; PF02401; LYTB; 1.
CC TIGRFAMs; TIGR00216; isph_lytB; 1.
CC IGTPREs; TIGR00216; Complete proteome; Oxidoreductase; NADP.
KW Isoprene biosynthesis;
SQ SEQUENCE 319 AA; 35741 MW; 69D8AFCC12DD09B8 CRC64;

Query Match 51.4%; Score 37; DB 1; Length 319;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SELGKSTNTF 12

```

```

DB 232 AELGKTGT 241
:|||||
RESULT 10
RPF2_HUMAN
ID RPF2_HUMAN STANDARD; PRT; 407 AA.
AC Q06858; Q9B047; Q9C021;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ret finger protein 2 (leukemia associated protein 5) (B-cell chronic
DE lymphocytic leukemia tumor suppressor Leu5) (Putative tumor suppressor
DE RFP2) (Tripartite motif protein 13).
GN RFP2 OR LEU5 OR TRIM13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Leukemia;
RX MEDLINE=98258969; PubMed=95390022;
RA Kapanadze B., Kasnuba V., Baranova A., Rasool O., van Everdink W.,
RA Liu Y., Symov A., Corcoran M., Poltarau A., Brodyansky V.,
RA Fedorova L., Kazakov A., Ibbotson R., van den Berg A., Gizatullin R.,
RA Grander D., Buys C., Oscier D., Zabarovsky E.R., Einhorn S.,
RA Yankovsky N.;
RA "A cosmid and cDNA fine physical map of a human chromosome 13q14
RA region frequently lost in B-cell chronic lymphocytic leukemia and
RA identification of a new putative tumor suppressor gene, Leu5.";
RL FEBS Lett. 426:266-270(1998).
[2]
SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=21100862; PubMed=11161783;
RA Kapanadze B., Makeeva N., Corcoran M., Jareborg N., Hammarsund M.,
RA Baranova A., Zabarovsky E., Vorontsova O., Merup M., Gahrton G.,
RA Jansson M., Yankovsky N., Einhorn S., Oscier D., Grander D.,
RA Sangfelt O.;
RA "Comparative sequence analysis of a region on human chromosome 13q14,
RA frequently deleted in B-cell chronic lymphocytic leukemia, and its
RA homologous region on mouse chromosome 14.";
RL Genomics 70:327-334(2000).
[3]
SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RX MEDLINE=21231161; PubMed=11331580;
RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,
RA Riganelli D., Zanaria E., Messali S., Cainarca S., Guffanti A.,
RA Minucci S., Pelicci P.G., Ballabio A.;
RA "The tripartite motif family identifies cell compartments.";
RL EMBO J. 20:2140-2151(2001).
[4]
SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=21164809; PubMed=11264177;
RA Migliazza A., Bosch F., Komatsu H., Cayanis E., Martinotti S.,
RA Toniato E., Guccione E., Qu X., Chien M., Murty V.V., Gaidano G.,
RA Inghirami G., Zhang P., Fischer S., Kalachikov S.M., Russo J.,
RA Edelman I., Efstratiadis A., Dalla-Favera R.;
RA "Nucleotide sequence, transcription map, and mutation analysis of the
RA 13q14 chromosomal region deleted in B-cell chronic lymphocytic
RA leukemia.";
RL Blood 97:2098-2104(2001).
[5]
SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

```

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May act as a tumor suppressor.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Alpha;
CC IsoId=O60858-1; Sequence=Displayed;
CC Name=Beta;
CC IsoId=O60858-2; Sequence=VSP_005746; VSP_005747;
CC -!- DISEASE: May be involved in B-cell chronic lymphocytic leukemia
CC (B-CLL) with a 13q14 region deletion.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ224819; CAA12136.1; -
CC EMBL: AF241849; AAK51624.1; -
CC EMBL: AF241850; AAF91315.1; -
CC EMBL: AF220127; AAG53500.1; -
CC EMBL: AF220128; AAG53501.1; -
CC EMBL: AF279660; AAK13059.1; -
CC EMBL: AL137060; CAC43391.1; -
CC EMBL: BC003579; AAH03579.1; -
CC Genew; HGNC:9976; RFP2.
CC MIM; 605661; -
CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
CC InterPro; IPR000315; Znf-Bbox.
CC InterPro; IPR001841; Znf_Ring.
CC Pfam; PF00643; zf-B_box; 1.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00336; BBOX; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00119; ZF_BBOX; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Anti-oncogene; Zinc-finger; Coiled coil; Alternative splicing;
CC Polymorphism.
CC ZN_FING 10 58 RING-TYPE
CC ZN_FING 89 131 B_BOX-TYPE
CC DOMAIN 172 200 COILED COIL (POTENTIAL).
CC VARSPLIC 175 175 L -> D (in isoform Beta).
CC VARSPLIC 176 407 /FTId=VSP_005746.
CC VARSPLIC 355 355 Missing (in isoform Beta).
CC VARSPLIC 355 355 /FTId=VSP_005747.
CC T -> S.
CC /FTId=VAR_013512.
CC SEQUENCE 407 AA; 47001 MW; E3B624345474AEBA CRC64;
SQ
Query Match 51.4%; Score 37; DB 1; Length 407;
Best Local Similarity 46.2%; Pred. No. 30;

```









CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z74817; CAA99085.1; -;  
 DR EMBL; Z74816; CAA99084.1; -;  
 DR PIR; S77690; S77690.  
 DR Germline; 143497; -;  
 DR SGD; S0005435; YOLO75C.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein;  
 KW Transport; Repeat.  
 FT TRANSMEM 376 396 POTENTIAL.  
 FT TRANSMEM 496 516 POTENTIAL.  
 FT TRANSMEM 531 551 POTENTIAL.  
 FT TRANSMEM 605 625 POTENTIAL.  
 FT TRANSMEM 1039 1059 POTENTIAL.  
 FT TRANSMEM 1121 1141 POTENTIAL.  
 FT TRANSMEM 1267 1287 POTENTIAL.  
 FT NP\_BIND 62 69 ATP (POTENTIAL).  
 FT NP\_BIND 727 734 ATP (POTENTIAL).  
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 983 983 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1294 AA; 145157 MW; C555500A45E9284E CRC64;

Query Match 51.4%; Score 37; DB 1; Length 1294;

Best Local Similarity 58.3%; Pred. No. 96;

Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CASELGKSTNTF 12

Db 1161 CGERLGIMTNTF 1172

# RESULT 15

UL16\_HCMVA  
 ID UL16\_HCMVA STANDARD; PRT; 230 AA.  
 AC P16757;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE Hypothetical protein UL16.  
 GN UL16.  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90269039; PubMed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
 RA Horenell T., Hutchison C.A. III, Kuzarides T., Martignetti J.A.,  
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
 RT "Analysis of the protein-coding content of the sequence of human  
 RT cytomegalovirus strain AD169."  
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X17403; CAA35448.1; -;

DR PIR; S09778; S09778.

KW Hypothetical protein.

FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 230 AA; 26147 MW; B72F2C241C569967 CRC64;

Query Match 50.7%; Score 36.5; DB 1; Length 230;

Best Local Similarity 57.1%; Pred. No. 20;

Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CASELG-KSTNTFC 13

Db 25 CAVDLGSKSSNTC 38

Search completed: September 5, 2004, 09:56:01

Job time : 7.90909 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 09:46:09 ; Search time 29.4141 Seconds  
(without alignments)  
139.448 Million cell updates/sec

Title: US-09-761-636A-5  
Perfect score: 72  
Sequence: 1 CASELGKSTNTFC 13

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	70.8	326	11 Q912E4	Q912E4 rattus norv
2	49	68.1	127	13 Q918V8	Q918V8 rana pipien
3	43	59.7	127	13 Q8UUV5	Q8UUV5 rana pipien
4	43	59.7	997	5 Q967Z1	Q967Z1 trypanosoma
5	42	58.3	1011	3 Q9P944	Q9P944 pneumocysti
6	41	56.9	316	13 Q902B2	Q902B2 brachydanio
7	40	55.6	283	16 Q7UWV5	Q7UWV5 rhodopirell
8	40	55.6	383	12 Q84551	Q84551 paramemum
9	40	55.6	605	16 Q81J08	Q81J08 bacillus an
10	40	55.6	605	16 Q81410	Q81410 bacillus ce
11	39.5	54.9	621	5 Q7YVH2	Q7YVH2 cryptospori
12	39	54.2	128	13 Q9DFV7	Q9DFV7 rana catesb
13	39	54.2	471	10 Q9FHK1	Q9FHK1 arabidopsis
14	39	54.2	677	10 Q9FLD9	Q9FLD9 arabidopsis
15	39	54.2	738	10 Q8L838	Q8L838 arabidopsis
16	39	54.2	1117	10 Q9M133	Q9M133 arabidopsis

17	39	54.2	2182	5 Q81I25	Q81I25 plasmodium
18	38	52.8	23	2 Q9R4A0	Q9R4A0 lactobacill
19	38	52.8	24	2 Q9R499	Q9R499 lactobacill
20	38	52.8	49	3 Q9UUP2	Q9UUP2 pneumocysti
21	38	52.8	128	13 Q9DFY8	Q9DFY8 rana catesb
22	38	52.8	254	15 Q89498	Q89498 human immun
23	38	52.8	254	15 Q9YXW6	Q9YXW6 human immun
24	38	52.8	255	15 Q89503	Q89503 human immun
25	38	52.8	258	3 P87104	P87104 pneumocysti
26	38	52.8	298	16 Q9S332	Q9S332 prochloroco
27	38	52.8	299	16 Q81TV5	Q81TV5 bacillus an
28	38	52.8	334	15 Q76115	Q76115 human immun
29	38	52.8	334	15 Q76113	Q76113 human immun
30	38	52.8	334	15 Q76104	Q76104 human immun
31	38	52.8	334	15 Q76107	Q76107 human immun
32	38	52.8	334	15 Q76088	Q76088 human immun
33	38	52.8	334	15 Q76098	Q76098 human immun
34	38	52.8	334	15 Q76106	Q76106 human immun
35	38	52.8	334	15 Q76108	Q76108 human immun
36	38	52.8	334	15 Q76084	Q76084 human immun
37	38	52.8	334	15 Q76083	Q76083 human immun
38	38	52.8	334	15 Q76080	Q76080 human immun
39	38	52.8	334	15 Q76109	Q76109 human immun
40	38	52.8	334	15 Q76116	Q76116 human immun
41	38	52.8	334	15 Q76101	Q76101 human immun
42	38	52.8	334	15 Q76110	Q76110 human immun
43	38	52.8	334	15 Q76094	Q76094 human immun
44	38	52.8	334	15 Q76114	Q76114 human immun
45	38	52.8	378	16 Q97GW2	Q97GW2 clostridium

## ALIGNMENTS

RESULT 1  
Q912E4 PRELIMINARY; PRT; 326 AA.  
ID Q912E4  
AC Q912E4;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE VEGF-D.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RA Kirkin V., Maitchev R., Krishnan J., Steffen A., Waltenberger J.,  
RA Pepper M.S., Giannis A., Sileman J.P.;  
RT "Characterization of indolinones which specifically inhibit VEGF-C and  
RT VEGF-D-induced activation of VEGFR-3 but not VEGFR-2";  
RL Eur. J. Biochem. 0:0-0(2001).  
DR EMBL; AV032728; AAK96008.1;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008083; F:growth factor activity; IEA.  
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.  
DR InterPro; IPR004153; CXXC repeat.  
DR InterPro; IPR000072; PD\_growth\_factor.  
DR Pfam; PF03128; CXXC; 1.  
DR Pfam; PF03411; PDGF; 1.  
DR ProDom; PD001629; PD\_growth\_factor; 1.  
DR SMART; SM00141; PDGF; 1.  
DR PROSITE; PS00249; PDGF\_1; 1.  
DR PROSITE; PS02278; PDGF\_2; 1.  
SQ SEQUENCE 326 AA; 37106 MW; D7CAEBA6C9FABB7D CRC64;

Query Match 70.8%; Score 51; DB 11; Length 326;  
Best Local Similarity 90.9%; Pred. No. 0.29;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ASELGKSTNTF 12

RP SEQUENCE FROM N.A.  
RX MEDLINE=20184731; PubMed=10721706;

RA Lee L.H., Gigliotti F., Wright T.W., Simpson-Haidaris P.J.,  
 RA Weinberg G.A., Haidaris C.G.;  
 RT "Molecular characterization of KBX1, a kexin-like protease in mouse  
 RT Pneumocystis carinii.";  
 RL Gene 242:141-150(2000);  
 CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.  
 DR EMBL: AF093132; AAF32493.1; -;  
 DR MEROPS; S08.011; -;  
 DR GO; GO:0008233; F:Peptidase activity; IEA.  
 DR GO; GO:0004289; F:Subtilase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR002029; Peptidase\_S8.  
 DR InterPro; IPR002884; Peptidase\_S8.  
 DR InterPro; IPR006970; PT.  
 DR InterPro; IPR002965; P rich extensn.  
 DR Pfam; PF00082; Peptidase\_S8; 1.  
 DR Pfam; PF04886; PT; 4.  
 DR Pfam; PF01483; P\_protein; 1.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR ProDom; PD000717; P domain; 1.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 KW Protease.  
 SQ SEQUENCE 1011 AA; 112021 MW; FBE472C8F65864E8 CRC64;  
 Query Match 58.3%; Score 42; DB 3; Length 1011;  
 Best Local Similarity 53.8%; Pred. No. 48;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CASELKGKSTNTFC 13  
 DB 246 CAGEIVAANKNTFC 258  
 |||:|:|  
 RESULT 6  
 Q90ZB2 PRELIMINARY; PRT; 316 AA.  
 AC Q90ZB2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Transcription factor Gbx1.  
 GN GBX1.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lun K., Thisse C., Thisse B., Amores A., Yan Y., Postlethwait J.,  
 RA Brand M.;  
 RT "Gbx2 but not gbx1 is dependent on fgf8 function.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDJJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; AF288763; AAK83070.1; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; Homeobox; 1.  
 DR PRINTS; PR00024; HOMEOBOX.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEOBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEOBOX\_2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 316 AA; 34408 MW; DC32955EA5430D8A CRC64;  
 Query Match 56.9%; Score 41; DB 13; Length 316;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFC 13  
 DB 63 ASFAGRLTNTFC 74  
 |||:|:|  
 RESULT 7  
 Q7UWW5 PRELIMINARY; PRT; 283 AA.  
 AC Q7UWW5;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN RB1746.  
 OS Rhodospirillum rubrum.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294135; CAD72247.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 283 AA; 31166 MW; 82926CDC0DF3A51A CRC64;  
 Query Match 55.6%; Score 40; DB 16; Length 283;  
 Best Local Similarity 58.3%; Pred. No. 30;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFC 13  
 DB 209 AABLGLSINSYC 220  
 |||:|:|  
 RESULT 8  
 Q84551 PRELIMINARY; PRT; 383 AA.  
 AC Q84551;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE A231L protein.  
 GN A231L.  
 OS Paramecium bursaria chlorella virus 1 (PBCV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
 OX NCBI\_TaxID=10506;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96187795; PubMed=8614977;  
 RA Lu Z., Li Y., Que Q., Kutish G.F., Rock D.L., Van Etten J.L.;  
 RT "Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map  
 RT positions 88 to 182.";  
 RL Virology 216:102-123(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20013326; PubMed=10544099;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., King W.,  
 RA Lisek A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chlorella virus PBCV-1 encodes a functional homosperridine  
 RT synthase.";  
 RL Virology 263:254-262(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20478054; PubMed=11021991;  
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus

```

RT PBCV-1.";
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42580; AAC96599.1; -.
DR FIR; T17722; T17722.
SQ SEQUENCE 383 AA; 43643 MW; B04E602D99973DCFC CRC64;

Query Match 55.6%; Score 40; DB 12; Length 383;
Best Local Similarity 46.2%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CASELKGKSTNTFC 13
Db |::|||
304 CVNIGKKKEFC 316

RESULT 9
Q81J08 PRELIMINARY; PRT; 605 AA.
AC Q81J08;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oligoendopeptidase F, putative.
GN BA5706.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Ostad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.W., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaitte J.E., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017041; AAP29338.1; -.
DR TIGR; BA5706; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

Query Match 55.6%; Score 40; DB 16; Length 605;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFC 13
Db |::|||
358 ADVNGKSTGAF 369

RESULT 10
Q81410 PRELIMINARY; PRT; 605 AA.
AC Q81410;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oligoendopeptidase F (EC 3.4.24.-).
GN BC5453.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapural V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis."
RL Nature 423:87-91(2003).
DR EMBL; AE017015; AAP12314.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001567; Peptidase_M3.
DR InterPro; IPR006025; Pept_M3n_BS.
DR Pfam; PF01432; Peptidase_M3; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 605 AA; 69123 MW; 3C16DB0B1B2F4062 CRC64;

Query Match 55.6%; Score 40; DB 16; Length 605;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFC 13
Db |::|||
358 ADVNGKSTGAF 369

RESULT 11
Q7YYH2 PRELIMINARY; PRT; 621 AA.
AC Q7YYH2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F-box domain protein.
GN IMB_228.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Bimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=Iowa;
RA  Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,
RA  Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
RT  "Integrated mapping, chromosomal sequencing and sequence analysis of
RT  Cryptosporidium parvum.";
RL  Genome Res. 0:0-0(2003).
DR  EMBL; BX538351; CAD98511.1; -.
SQ  SEQUENCE 621 AA; 71995 MW; D99793733B5A3658 CRC64;

Query Match      54.9%; Score 39.5; DB 5; Length 621;
Best Local Similarity 56.2%; Pred. No. 86;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY  1 CASEL---GKSTNTFC 13
Db  161 CASESTFEGERQNTFC 176

RESULT 12
Q9DFY7
ID  Q9DFY7 PRELIMINARY; PRT; 128 AA.
AC  Q9DFY7
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  RC-RNase3 ribonuclease precursor.
OS  Rana catesbeiana (Bull frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX  NCBI_TaxID=8400;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RX  MEDLINE=20512555; PubMed=11058105;
RA  Liao Y.D., Huang H.C., Liu Y.J., Wei C.W., Tang P.C., Wang S.C.;
RT  "Purification and cloning of cytotoxic ribonucleases from Rana
RT  catesbeiana (bullfrog).";
RL  Nucleic Acids Res. 28:4097-4104(2000).
DR  EMBL; AF242554; AAG31440.2; -.
DR  HSSP; P22069; IONC.
DR  GO; GO:0003676; F:nucleic acid binding; IEA.
DR  GO; GO:0004522; F:pancreatic ribonuclease activity; IEA.
DR  InterPro; IPR001427; RNaseA.
DR  Pfam; PF00074; rnasea; 1.
DR  ProDom; PD000535; RNaseA; 1.
DR  PROSITE; PS00127; RNASE_PANCREATIC; 1.
KW  Signal.
FT  SIGNAL 1 23 POTENTIAL.
FT  CHAIN 24 128 RC-RNASE3 RIBONUCLEASE.
SQ  SEQUENCE 128 AA; 14517 MW; 2B14986082E0587D CRC64;

Query Match      54.2%; Score 39; DB 13; Length 128;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY  1 CASELGSSTNTFC 13
Db  98 CHYKLSSTNTFC 110

RESULT 13
Q9FHK1
ID  Q9FHK1 PRELIMINARY; PRT; 471 AA.
AC  Q9FHK1
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Genomic DNA, chromosome 5, p1 clone:MAB16.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

```

```

OC  eurosids II; Brassicales; Brassicaceae; Arabidopsais.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Columbia;
RX  MEDLINE=20181125; PubMed=10718197;
RA  Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA  Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT  features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT  clones.";
RL  DNA Res. 7:31-63(2000).
DR  EMBL; AB018112; BAB10975.1; -.
DR  EMBL; AB026661; BAB10975.1; JOINED.
DR  InterPro; IPR001810; F-box.
DR  InterPro; IPR006527; F-box assoc 1.
DR  Pfam; PF00646; F-box; 1.
DR  SMART; SM00256; FBOX; 1.
DR  TIGRFAMs; TIGR01640; F_box_assoc_1; 1.
DR  PROSITE; PS0181; FBOX; 1.
SQ  SEQUENCE 471 AA; 55212 MW; 1E5D997B7942E26C CRC64;

Query Match      54.2%; Score 39; DB 10; Length 471;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY  4 ELGKSTNTFC 13
Db  365 DLGESINKYC 374

RESULT 14
Q9FLD9
ID  Q9FLD9 PRELIMINARY; PRT; 677 AA.
AC  Q9FLD9
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE  Similarity to hedgehog-interacting protein.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsais.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Columbia;
RX  MEDLINE=98290546; PubMed=9628582;
RA  Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA  Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT  Sequence features of the regions of 1,456,315 bp covered by nineteen
RT  physically assigned p1 and TAC clones.";
RL  DNA Res. 5:41-54(1998).
DR  EMBL; AB010077; BAB10221.1; -.
SQ  SEQUENCE 677 AA; 74196 MW; F0C2B09BBE1B079 CRC64;

Query Match      54.2%; Score 39; DB 10; Length 677;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY  3 SELGKSTNTFC 13
Db  144 SEIWKSSNDFC 154

RESULT 15
Q8L838
ID  Q8L838 PRELIMINARY; PRT; 738 AA.
AC  Q8L838
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
GN AT4G01400.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sekano H., Kawai J.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY120763; AM53321.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 738 AA; 82977 MW; EF9B1B0FDA7B15CE CRC64;  
  
Query Match 54.2%; Score 39; DB 10; Length 738;  
Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CASELGKSTNTF 12  
| |||||: :||  
Db 543 CLSELGELSSTF 554

Search completed: September 5, 2004, 09:59:52  
Job time : 31.4141 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2004, 09:37:49 ; Search time 42.6768 Seconds  
(without alignments)  
86.068 Million cell updates/sec

Title: US-09-761-636A-5  
Perfect score: 72  
Sequence: 1 CASELGKSTNTFC 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	72	100.0	13	4	AAU04524
2	72	100.0	17	4	AAU04527
3	69	95.8	13	4	AAU04534
4	69	95.8	17	4	AAU04550
5	69	95.8	17	4	AAU04546
6	65	90.3	13	4	AAU04535
7	65	90.3	13	4	AAU04537
8	65	90.3	17	4	AAU04547
9	65	90.3	17	4	AAU04551
10	65	90.3	17	4	AAU04549
11	61	84.7	13	4	AAU04536
12	61	84.7	17	4	AAU04548
13	54	75.0	73	4	AAU04522
14	54	75.0	96	4	AAU04520
15	54	75.0	109	2	AAU23889
16	54	75.0	109	3	AAB11931
17	54	75.0	109	6	ABB84621
18	54	75.0	109	6	ABG73750
19	54	75.0	287	6	ABG73779
20	54	75.0	325	2	AAW53240
21	54	75.0	325	4	AAU04522
22	54	75.0	354	2	AAU44293
23	54	75.0	354	2	AAW49036
24	54	75.0	354	2	AAW53241
25	54	75.0	354	3	AAB10649

26	54	75.0	354	3	AAU04524
27	54	75.0	354	3	AAU04527
28	54	75.0	354	3	AAU04534
29	54	75.0	354	4	AAU04550
30	54	75.0	354	4	AAU04546
31	54	75.0	354	4	AAU04535
32	54	75.0	354	4	AAU04537
33	54	75.0	354	4	AAU04551
34	54	75.0	354	5	ABG73750
35	54	75.0	354	6	ABG73779
36	54	75.0	354	7	AAW53240
37	54	75.0	620	2	AAU04522
38	51	70.8	110	5	AAU04520
39	51	70.8	178	2	AAU04522
40	51	70.8	321	2	AAU04522
41	51	70.8	321	5	AAU04522
42	51	70.8	326	2	AAU04522
43	51	70.8	337	2	AAU04522
44	51	70.8	358	2	AAU04522
45	51	70.8	358	2	AAU04522

ALIGNMENTS

RESULT 1  
AAU04524  
ID AAU04524 standard; peptide; 13 AA.  
XX AC AAU04524;  
XX AC  
DT 26-SEP-2001 (first entry)  
XX DE VEGF based monocyclic peptide 1.  
XX KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
KW diabetic retinopathy; chronic inflammation; cyclic.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Disulfide-bond 1..13 /note= "This bond cyclises the peptide"  
XX PN WO200152875-A1.  
XX PD 26-JUL-2001.  
XX PF 18-JAN-2001; 2001WO-US001533.  
XX PR 18-JAN-2000; 2000US-0176293P.  
XX PR 16-MAY-2000; 2000US-0204590P.  
XX PA (LUDW-) LUDWIG INST CANCER RES.  
XX PI Achen MG, Hughes RA, Stacker S, Cendron A;  
XX WP1; 2001-442248/47.  
XX PT Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
XX PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
XX PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
XX PT residues.  
XX PS Claim 49; Page 32; 102pp; English.  
XX CC The sequence represents a monomeric monocyclic peptide of the invention,  
XX CC whose 3-dimensional structure is modelled on the exposed loop of human  
XX CC VEGF (vascular endothelial growth factor). The invention relates to a  
XX CC method of producing a monomeric monocyclic peptide by a measuring beta-  
XX CC beta carbon separation distances on opposite antiparallel strands of a

CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 72; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0;

QY 1 CASELGKSTNTFC 13  
 |||||  
 Db 1 CASELGKSTNTFC 13

RESULT 2

AAU04527  
 ID AAU04527 standard; protein; 17 AA.

AC AAU04527;

XX 26-SEP-2001 (first entry)

XX VEGF based bicyclic dimeric peptide #1.

XX Human; VEGF; vascular endothelial growth factor; angiogenesis;  
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..13 /note= "This bond cyclises the peptide"

FT Disulfide-bond 17

FT /note= "A disulfide bond forms between residue 17 and  
 residue 17 of an identical peptide to form a dimeric  
 peptide, or to residue 1 of the sequence appearing as  
 AAU04528, also forming a dimeric peptide"

XX WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US001533.

XX 18-JAN-2000; 2000US-0176293P.

XX 16-MAY-2000; 2000US-0204590P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX

DR WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
 PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
 PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
 PT residues.

XX Claim 59; Page 32; 102pp; English.

XX The sequence represents a dimeric bicyclic peptide of the invention,  
 CC whose 3-dimensional structure is modelled on the expose loop of human  
 CC VEGFD (vascular endothelial growth factor). The invention relates to a  
 CC method of producing a monomeric monocyclic peptide by a measuring beta-  
 CC beta carbon separation distances on opposite antiparallel strands of a  
 CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX

SQ Sequence 17 AA;

Query Match 100.0%; Score 72; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 2.8e-05; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0;

QY 1 CASELGKSTNTFC 13  
 |||||  
 Db 1 CASELGKSTNTFC 13

RESULT 3

AAU04534

ID AAU04534 standard; peptide; 13 AA.

XX AAU04534;

XX 26-SEP-2001 (first entry)

XX VEGF based monocyclic peptide 12.

XX Human; VEGF; vascular endothelial growth factor; angiogenesis;  
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..13 /note= "This bond cyclises the peptide"

FT

XX WO200152875-A1.

XX 26-JUL-2001.

XX

PF 18-JAN-2001; 2001WO-US001533.  
 XX  
 PR 18-JAN-2000; 2000US-0176293P.  
 PR 16-MAY-2000; 2000US-0204590P.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Achen MG, Hughes RA, Stacker S, Cendron A;  
 XX  
 DR WPI; 2001-442248/47.  
 XX  
 XX Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
 PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
 PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
 PT residues.  
 XX  
 XX Example 25; Page 47; 102pp; English.  
 XX  
 CC The sequence represents a monomeric monocyclic peptide of the invention,  
 CC whose 3-dimensional structure is modelled on the expose loop of human  
 CC VEGF (vascular endothelial growth factor). The invention relates to a  
 CC method of producing a monomeric monocyclic peptide by a measuring beta-  
 CC beta carbon separation distances on opposite antiparallel strands of a  
 CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 13 AA;  
 Query Match 95.8%; Score 69; DB 4; Length 13;  
 Best Local Similarity 92.3%; Pred. No. 6.8e-05;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 CASELGKSTNTFC 13  
 Db ||:|||||:|  
 1 CATELGKSTNTFC 13  
 RESULT 4  
 ID AAU04550 standard; peptide; 17 AA.  
 XX  
 AC AAU04550;  
 XX  
 XX 26-SEP-2001 (first entry)  
 DE VEGF based bicyclic dimeric peptide #7.  
 XX  
 KW Human; VEGF; vascular endothelial growth factor; angiogenesis;  
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;  
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;  
 KW diabetic retinopathy; chronic inflammation; cyclic.  
 XX  
 OS Synthetic.

XX  
 PH Key Location/Qualifiers  
 FT Disulfide-bond 1. 13  
 FT /note= "This bond cyclises the peptide"  
 FT Disulfide-bond 17  
 FT /note= "A disulfide bond forms between residue 17 and  
 FT residue 1 of the sequence appearing as AAU04528, forming  
 FT a dimeric peptide"  
 XX  
 PN WO200152875-A1.  
 XX  
 XX 26-JUL-2001.  
 PD  
 PD 18-JAN-2001; 2001WO-US001533.  
 PF  
 PF 18-JAN-2000; 2000US-0176293P.  
 PR 16-MAY-2000; 2000US-0204590P.  
 XX  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA  
 PI Achen MG, Hughes RA, Stacker S, Cendron A;  
 XX  
 DR WPI; 2001-442248/47.  
 XX  
 XX Novel monomeric monocyclic peptide, used to interfere with angiogenesis,  
 PT or lymphangiogenesis, is produced by cyclizing a peptide loop fragment  
 PT from an exposed loop of a growth factor protein by oxidizing the cysteine  
 PT residues.  
 XX  
 XX Example 26; Page 49; 102pp; English.  
 PS  
 CC The sequence represents a dimeric bicyclic peptide of the invention,  
 CC whose 3-dimensional structure is modelled on the expose loop of human  
 CC VEGF (vascular endothelial growth factor). The invention relates to a  
 CC method of producing a monomeric monocyclic peptide by a measuring beta-  
 CC beta carbon separation distances on opposite antiparallel strands of a  
 CC peptide loop fragment from an exposed loop of a growth factor protein and  
 CC cyclising the peptide by oxidising the cysteine residues. The monocyclic  
 CC peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic  
 CC peptides) and a cyclic peptide with at least one amino acid deleted prior  
 CC to cyclisation are used to interfere with angiogenesis,  
 CC neovascularisation or lymphangiogenesis in a mammal with a condition  
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.  
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,  
 CC hemangioma, vascularised malignant or benign tumour, post-recovery  
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold  
 CC trauma, substance-induced neovascularisation of the liver, excessive  
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular  
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver  
 CC infection. The peptides are also used to modulate vascular permeability  
 CC in a mammal (the mammal has a condition characterised by fluid  
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,  
 CC or brain. The peptides are used to image blood vessels and lymphatic  
 CC vasculature. The monomeric and bicyclic peptides are used to interfere  
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and  
 CC are also used in combination with an anti-inflammatory agent, to treat a  
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and  
 CC diabetic retinopathy  
 XX  
 SQ Sequence 17 AA;  
 Query Match 95.8%; Score 69; DB 4; Length 17;  
 Best Local Similarity 92.3%; Pred. No. 9.1e-05;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 CASELGKSTNTFC 13  
 Db ||:|||||:|  
 1 CASELGKSTNTFC 13  
 RESULT 5  
 ID AAU04546 standard; peptide; 17 AA.

